Appunti di Analisi Matematica

Mattia Belletti

5 luglio 2004

Indice

Capitolo 1

Licenza

Segue il testo originale della GFDL (GNU Free Document License) sotto cui è distribuito questo documento:

Una sua traduzione è (approssimativamente ;-)):

Questo prodotto non è stato compilato e fornito dall'autore, una copia originale e' reperibile in http://www.cs.unibo.it/ mbellett/

le differenze sono colpa di Matteo Boccafoli 2003

Capitolo 2

Introduzione

Questo testo é una versione ordinata degli appunti del corso di Analisi Matematica da me presi durante le lezioni del professore C. Parenti ed integrati dal testo Matematica di M. Bramanti, C. Pagani, S. Salsa a cui lo stesso professore ha fatto esplicito riferimento durante le sue lezioni.

Per ogni correzione o chiarimento riguardo a questo testo, sono reperibile per mezzo dell'indirizzo mail mbellett@cs.unibo.it - non é gradito nessun genere di spam!

Questo documento è nato per varie ragioni: la prima, è stata, lo ammetto, quella di aiutare *me*, per imparare meglio la materia, ragionarci sopra, "tappare" i buchi che potevano esser stati lasciati a lezione per mancanza di tempo, e ripetere. Inoltre, non poco mi è servita per imparare l'utilizzo di Latex (una serie di macro per Tex, a sua volta un programma nato per la formattazione di testo, utilizzato anche a livello professionale ed estremamente immediato). Infine, un pochino volevo anche aiutare gli altri! :).

Ringraziamenti? No, non sono solito ringraziare, perchè ho la brutta abitudine di dimenticare chi se lo merita e ricordare chi lo merita di meno. Chi ha la mia gratitudine, lo sa già, senza bisogno che ci sia questo documento a ricordarlo al mondo :).

Per chi notasse differenze di versione rispetto all'originale: sì, ho deciso di ri-numerare le versioni (nella quasi certezza di non causare molti casini :P), per riflettere la non-completezza degli appunti.
Versione: 0.9-pre1.

History:

- Versione 0.9-pre1 (03 Luglio 2004): nessun cambiamento di contenuti decisivo da parte mia, qualche cambiamento da parte di Matteo Boccafoli - però questa è la prima release su appforge :).
- Versione 0.3 (26 Settembre 2002): riguardati i capitoli fino al terzo semi-incluso :), aggiunte un paio di immagini, in programma di inserire vari grafici di funzione in giro con l'aiuto di gnuplot.
- Versione 0.2 (3 Luglio 2002): aggiornata la prefazione, inserita come licenza la GFDL (GNU Free Documentation License, per la sua stesura completa, vedere il capitolo precedente), cambiato lo stile da `article` a `book`.
- Versione 0.1 (intorno al Novembre 2001): mantenuta dalla nascita del documento fino ad una sua quasi completezza: assente ancora il finale, la trasformazione di coordinate per gli integrali multipli. E, soprattutto, manca una licenza valida!

Ultima revisione: 5 luglio 2004

Capitolo 3

Veloci richiami ai concetti basilari

3.1 Insiemi

Il concetto di *insieme* è un concetto 'primitivo' del linguaggio matematico, che cioè non è definibile utilizzando concetti più elementari. Intuitivamente, il formalismo per esso utilizzato rappresenta il concetto di una "collezione", "raccolta" di oggetti unici (nel senso che non sono presenti due oggetti a e b "duplicati", ovvero per cui valga $a = b$, ma che allo stesso tempo vengano considerati distinti nell'insieme).

La notazione solitamente utilizzata è questa: con le lettere *maiuscole* $(A, B, C, \dots)$ si indicano gli insiemi, con le lettere *minuscole* $(a, b, c, \dots)$ gli elementi che fanno parte di questi insiemi.

Il simbolo

$$\in$$

indica l'appartenenza di un oggetto ad un insieme. Ad esempio

$$a \in A$$

significa che l'elemento a è un oggetto (o *elemento*) dell'insieme A.

Se si vuole indicare un elemento elencandone gli elementi, si utilizzano le parentesi graffe. Ad esempio

$$A = \{\, a, b, c \,\}$$

indica che l'insieme A è composto dai tre elementi a, b e c.

Due insiemi sono uguali se e solo se hanno gli stessi elementi; ciò viene indicato con il semplice simbolo di uguaglianza:

$$A = B$$

Si dice invece che B è un *sottoinsieme* di A se ogni elemento di B è un elemento di A, e si scrive:

$$B \subseteq A$$

oppure

$$A \supseteq B$$

Questa notazione non esclude tuttavia che $A = B$. Nel caso ciò non possa essere vero (qualche elemento di A non è elemento di B) si scrive:

$$B \subset A$$

oppure

$$A \subset B$$

In questo caso si dice che B è *sottoinsieme proprio* di A.

Bisogna infine anche ricordare un insieme particolare, quello che non contiene nessun elemento, chiamato *insieme vuoto* e che si indica con $\emptyset$.

Le relazioni di inclusione $\subseteq$ e $\supseteq$ sono *relazioni d'ordine*, ovvero godono delle proprietà *riflessiva* ($A \subseteq A$), *antisimmetrica* ($A \subseteq B \ \wedge \ B \subseteq A \implies A = B$) e *transitiva* ($A \subseteq B \ \wedge \ B \subseteq C \implies A \subseteq C$).

Sono definite alcune operazioni tra insiemi: l'*unione* di due insiemi A e B si scrive $A \cup B$ ed indica l'insieme di tutti quegli elementi che appartengono ad almeno uno dei due insiemi; l'*intersezione* di due insiemi A e B si scrive $A \cap B$ ed indica l'insieme di tutti quegli elementi che appartengono ad entrambi gli insiemi. Se $A \cap B = \emptyset$, si dice che A e B sono *disgiunti*. La differenza $A \setminus B$ è l'insieme costituito da tutti gli elementi di A che *non* appartengono a B.

L'insieme di tutti i sottoinsiemi propri ed impropri di U si chiama *insieme delle parti* di U e si indica con $\mathcal{P}(U)$. Se $A \in \mathcal{P}(U)$, si definisce *complementare* di A rispetto ad U (chiamato in questo caso *insieme ambiente* o *insieme universo*), e si indica con $\complement_U A$ (o semplicemente $\complement A$ se l'ambiente è chiaro dal contesto) l'insieme formato dagli elementi di U che non appartengono ad A, ovvero $\complement_U A = U \setminus A$.

Viene chiamato *prodotto cartesiano* di due insiemi A e B, ed indicato con $A \times B$, l'insieme costituito da tutte le possibili coppie ordinate (a, b) prodotte prendendo $a \in A$ e $b \in B$. Con "coppia ordinata" si intende una coppia di elementi in cui l'ordine è importante – per cui, quindi, $(a, b) \neq (b, a)$.

3.2 Sommatorie

La scrittura

$$\sum_{i=m}^{n} a_i$$

viene chiamata *sommatoria* per i che va da m ad n degli a_i, ed indica in modo conciso la somma:

$$a_m + a_{m+1} + a_{m+2} + \cdots + a_{n-1} + a_n$$

i viene chiamato *indice della sommatoria* ed è un indice *muto* - ovvero, cambiando nell'espressione tutte le occorrenze del simbolo con un altro non già usato, il significato non cambia. Questo non vale per n ad esempio (e ciò è di facile verifica).

Le proprietà delle sommatorie sono tutte abbastanza intuitive:

1. Prodotto per una costante:
$$\sum_{k=m}^{n} (c \cdot a_k) = c \cdot \sum_{k=m}^{n} a_k$$
2. Sommatoria di termini costanti:
$$\sum_{k=m}^{n} c = c \cdot (n - m + 1)$$
3. Somma di sommatorie:
$$\sum_{k=m}^{n} a_k + \sum_{k=m}^{n} b_k = \sum_{k=m}^{n} (a_k + b_k)$$
4. Composizione di sommatorie:
$$\sum_{k=m}^{n} a_k + \sum_{k=n+1}^{n+h} a_k = \sum_{k=m}^{n+h} a_k$$
5. Traslazione degli indici:
$$\sum_{k=m}^{n} a_k = \sum_{k=m+h}^{n+h} a_{k-h}$$
6. Riflessione degli indici:
$$\sum_{k=m}^{n} a_k = \sum_{k=m}^{n} a_{n+m-k}$$

3.3 Fattoriale e coefficienti binomiali

Il fattoriale del numero naturale n si indica con $n!$ ed indica il prodotto dei primi n numeri interi maggiori di 0, ovvero:

$$n! = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot (n-1) \cdot n$$

Per definizione, si pone $0! = 1$. Alcune semplici proprietà del fattoriale sono:

$$n! = n(n-1)!$$

Che insieme alla regola, posta per definizione, $0! = 1$, indica anche un modo alternativo per definire questa operazione in una formula ricorsiva:

$$n! = \begin{cases} 1, & \text{se } n = 0 \\ n(n-1)!, & \text{se } n > 0 \end{cases}$$

È inoltre utile ricordare anche la seguente proprietà

$$\text{se } 1 < k < n \quad \frac{n!}{(n-k)!} = n(n-1)(n-2)\ldots(n-k+1)$$

Un concetto che fa uso del fattoriale è quello dell'operatore *coefficiente binomiale* - la sua notazione e definizione sono le seguenti:

$$\text{se } 0 \leq k \leq n \quad \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

che, per quanto osservato prima, è uguale a:

$$\binom{n}{k} = \frac{n(n-1)(n-2)\ldots(n-k+1)}{k!}$$

Di solito si utilizza direttamente questa notazione, in quanto permette di definire il coefficiente binomiale non solo che se n e k sono interi, ma anche quando n è un qualunque numero reale (vedi anche 3.4 e 3.6). Un binomiale inteso con primo termine anche reale viene chiamato *coefficiente binomiale generalizzato* (6.10).

La cosiddetta formula di Newton utilizza i coefficienti binomiali per definire la potenza n-esima di una somma:

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

Infine, sono da ricordare anche le seguenti proprietà dei binomiali:

$$\binom{n}{k} = \binom{n}{n-k}$$

e

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$$

3.4 Naturali, Interi, Razionali

Gli insiemi su cui più spesso ci si trova a lavorare sono gli insiemi numerici.

Il primo insieme numerico che si incontra e si studia è quello dei numeri naturali ($\mathbb{N}$), ovvero tutti i numeri interi positivi e lo 0:

$$\mathbb{N} = \{\, 0, 1, 2, \dots \}$$

Un'estensione di questo insieme di numeri sono i numeri interi ($\mathbb{Z}$) ovvero $\mathbb{N}$ con l'aggiunta di anche i numeri negativi. Risulta quindi che $\mathbb{N}$ è un sottoinsieme di $\mathbb{Z}$:

$$\mathbb{Z} = \{\, 0, 1, -1, 2, -2, \dots \}$$

In seguito si trovano i numeri razionali ($\mathbb{Q}$), ovvero tutte le frazioni con il numeratore appartenente a $\mathbb{Z}$ e il denominatore appartenente a $\mathbb{Z}$ escluso lo 0:

$$\mathbb{Q} = \left\{ \frac{p}{q} \mid p, q \in \mathbb{Z},\ q \neq 0 \right\}$$

Quindi $2/3$, $-7/5$, $-9/-3$ sono elementi di $\mathbb{Q}$, mentre $2/0$ non lo è.

Ora, se $a \in \mathbb{Z}$, allora si può far corrispondere ad a l'elemento corrispondente di $\mathbb{Q}$: $a/1$. Se si accetta ciò, si può considerare valido il fatto che $\mathbb{Z} \subset \mathbb{Q}$.

3.5 Proprietà di $\mathbb{Q}$

$\mathbb{Q}$ è un'insieme strutturato, e in particolare possiede le operazioni $+$ e $\times$ (ovvero $(\mathbb{Q}, +, \times)$ è una struttura algebrica). Queste operazioni sono interne, ovvero:

$$a, b \in \mathbb{Q} \Rightarrow \begin{cases} a + b \in \mathbb{Q} \\ a \times b \in \mathbb{Q} \quad \text{(ovvero } ab \in \mathbb{Q}\text{)} \end{cases}$$

Inoltre possiedono varie altre proprietà. Entrambe sono associative, commutative ed possiedono elemento neutro (0 per l'addizione, 1 per la moltiplicazione). In simboli:

$$\forall a, b, c \in \mathbb{Q}\ : \quad \begin{array}{ll} a + (b + c) = (a + b) + c & a \cdot (b \cdot c) = (a \cdot b) \cdot c \\ a + b = b + a & a \cdot b = b \cdot a \\ a + 0 = 0 + a = a & a \cdot 1 = 1 \cdot a = a \end{array}$$

Inoltre, per quanto riguarda l'addizione, è vero che $\forall a, \exists b : a + b = 0$, (da ricordare che 0 è l'elemento neutro), ovvero per ogni a, esiste il suo opposto. Per quanto riguarda la moltiplicazione, anche in questo caso per qualsiasi a esiste il suo opposto, ma solo se $a \neq 0$.

Inoltre vale la proprietà distributiva che collega le due operazioni: $a(b + c) = ab + ac$.

$\mathbb{Q}$ è poi ordinato, ovvero esiste una relazione ($\leq$) di confronto. Questa proprietà gode di alcune proprietà:

$$\begin{array}{rcl} \forall a, b \in \mathbb{Q}: & a \leq b \vee b \leq a & \text{(proprietà del terzo escluso)} \\ \forall a \in \mathbb{Q}: & a \leq a & \text{(proprietà riflessiva)} \\ \forall a, b, c \in \mathbb{Q}: & a \leq b \wedge b \leq c \Rightarrow a \leq c & \text{(proprietà transitiva)} \\ \forall a, b \in \mathbb{Q}: & a \leq b \wedge b \leq a \Rightarrow a = b & \text{(proprietà antisimmetrica)} \end{array}$$

È poi facile definire le altre relazioni di confronto a partire da questa:

$$a < b \equiv a \leq b \wedge a \neq b$$
$$a > b \equiv \neg(a \leq b)$$
$$a \geq b \equiv a > b \vee a = b$$

C'è anche da considerare che $\leq, +$ e $\times$ sono legati da due principali proprietà.

$$\forall a, b, c \in \mathbb{Q} : a \leq b \Rightarrow a + c \leq b + c \tag{3.1}$$
$$\forall a, b, c \in \mathbb{Q}, 0 \leq c : ac \leq bc$$
$$\forall a, b, c \in \mathbb{Q}, c \leq 0 : bc \leq ac \tag{3.2}$$

Dalla 3.1 si può inoltre facilmente ricavare la relazione contraria:

$a + c \leq b + c \Rightarrow$
$a + c - c \leq b + c - c$ (ho aggiunto ad entrambi i membri -c)
$a + (c - c) \leq b + (c - c)$
$a + 0 \leq b + 0$
$a \leq b$

Come si può vedere quindi, conseguenza della prima proprietà è anche la sua inversa - è quindi giusto scrivere $\forall a, b, c \in \mathbb{Q} : a \leq b \Leftrightarrow a + c \leq b + c$.

3.5.1 Rappresentazioni dei numeri razionali

I numeri di $\mathbb{Q}$ possono anche essere rappresentati come *stringhe* decimali, dove con "stringa" intendiamo una serie di simboli. Nel caso dei numeri razionali, il tipo di stringa che si può costruire è questo:

$$\alpha, \beta_1\beta_2\beta_3 \ldots \beta_n \ldots$$

Dove $\alpha \in \mathbb{Z}$ e $\beta_i \in \mathbb{N}$. Tuttavia, com'è facile convincersi esaminando la semplice regola per la divisione di due numeri interi, questa stringa, per quanto riguarda i numeri razionali, deve rispettare certe regole. Essa può infatti assumere solo due forme: o $\exists n \in \mathbb{N} : \forall m \geq n : \beta_m = 0$, ovvero da un certo β_n in poi, vi sono solo 0; oppure il numero è periodico, cioè un gruppo di cifre nella sua parte decimale continua da un certo punto in poi a ripetersi sempre uguale.

Vi sono ovviamente dei metodi per passare dalla forma di frazione a quella di striga decimale. Quello per il passaggio diretto è dato dalla semplice regola della divisione. Quello per il procedimento inverso è invece più laborioso. Definiamo inanzitutto α come *parte intera* della stringa, e la sequenza dei β, finito o infinito che esso sia, come *parte decimale*. All'interno della parte decimale, è chiamato *antiperiodo* la parte precedente all'inizio del periodo (se esiste il periodo, altrimenti tutta la parte decimale è comunque impropriamente detta antiperiodo). Il *periodo* è invece, ovviamente, l'insieme di cifre che sempre si ripetono uguali. La regola per passare dalla stringa decimale alla frazione è quindi questa: si crea una frazione con, al numeratore, la differenza tra il numero formato da parte intera e parte decimale una consecutiva all'altra, e la parte intera, mentre al denominatore si pone un numero formato da tanti 9 quante sono le cifre del periodo, seguiti da tanti 0 quante sono le cifre dell'antiperiodo. La giustificazione di questo metodo sarà data più avanti, nello studio delle serie.

Esempio 3.5.1 *Convertire il numero razionale*

$$2,3747474 \cdots = 2,3\overline{74}$$

dalla sua rappresentazione decimale a quella frazionaria.

La soluzione risulta alquanto facile:

$$\frac{2374 - 23}{990} = \frac{2351}{990}$$

3.5.2 La retta numerica

Una rappresentazione standard dei numeri è quella sulla *retta numerica* (figura 3.1). Questa retta è definita da 2 soli punti: quello che rappresenta lo 0 e quello che rappresenta l'1. In questo modo viene anche definito il verso (quello che da 0 va a 1) e la scala utilizzata (nel caso serva). Su questa retta si possono ora rappresentare tutti i numeri razionali. Per di più, ogni punto della retta numerica definisce una lunghezza (quella tra sè e lo 0) che è commensurabile. In generale, due grandezze sono commensurabili tra loro se esiste un multiplo della prima ed un multiplo della seconda (il più delle volte secondo due interi diversi) che sono uguali. Ci si può facilmente convincere di questa proprietà tra i numeri razionali e l'unità.

Figura 3.1: La retta numerica

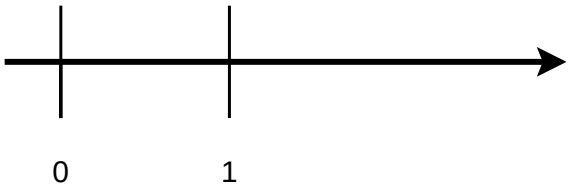
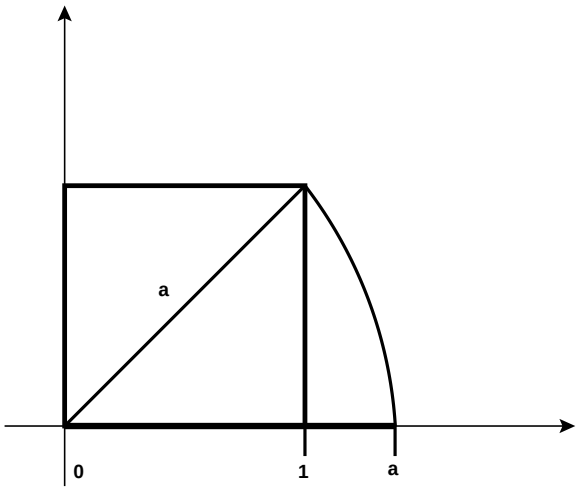


Figura 3.2: Un quadrato e la sua diagonale



3.6 Reali

Tuttavia non tutti i punti sulla retta corrispondono a dei numeri di $\mathbb{Q}$, e questo si può notare grazie a vari esempi. Uno dei più semplici è quello che segue.

Seguendo la costruzione riportata (figura 3.2), abbiamo identificato il punto a. Questo punto corrisponde ad un numero di $\mathbb{Q}$? La risposta, come vedremo, è negativa, e ciò si può rendere chiaro per mezzo di una dimostrazione per assurdo. Una dimostrazione per assurdo funziona in questo modo: per dimostrare l'affermazione $\neg P$, si parte col supporre invece P, cosa che sappiamo falsa ma che comunque accettiamo come vera. Se si riesce, ragionando su questa asserzione, ad arrivare ad una contraddizione, allora si deve accettare che l'ipotesi era errata, e quindi che sia vero il contrario: $\neg P$.

Quindi, partiamo col supporre che $a \in \mathbb{Q}$. Possiamo senz'altro anche supporre quindi che $a > 0$, in base al fatto che stiamo misurando una lunghezza e quindi un valore maggiore di 0. Inoltre, dal teorema di Pitagora sappiamo che $a \times a = 2$. Inoltre, dato che, per ipotesi, $a \in \mathbb{Q}$, allora, possiamo anche affermare che $\exists p, q \in \mathbb{N}, p, q > 0 : a = p/q$. Possiamo affermare che p e q sono naturali maggiori di zero perché la lunghezza del segmento è senz'altro maggiore di 0. Inoltre sappiamo anche che è sempre possibile trovare tra tutte le possibili coppie di p e q una per la quale questi due numeri siano primi tra loro, ovvero non abbiano divisori in comune.

A questo punto troviamo che $p^2/q^2 = a^2 = 2$. Quindi, $p^2 = 2q^2$, ovvero p^2 è un divisibile per 2 (pari). Ora, se p^2 è pari, lo è anche p - per arrivare a ciò basta analizzare la fattorizzazione di un numero: il fattore 2 può apparire nel suo quadrato solo se era presente anche nella base. Quindi, possiamo rappresentare p come $2r, r \in \mathbb{N}$. Sostituiamo quindi nella formula $p^2 = 2q^2$, che diventa così $4r^2 = 2q^2$, o meglio, semplificata, $2r^2 = q^2$. Seguendo un ragionamento analogo al precedente, dobbiamo concludere che anche q è un multiplo di 2. Ma questo vuol dire che p e q hanno dei fattori in comune: siamo caduti in contraddizione.

Ecco quindi che, seguendo lo schemo del ragionamento per assurdo, dobbiamo arrivare ad affermare che era errata l'ipotesi, e che quindi $a \notin \mathbb{Q}$. Questo risultato mostra quindi che poter rappresentare certi numeri occorre espandere l'insieme dei numeri. Così viene a formarsi l'insieme dei numeri reali $\mathbb{R}$. Possiamo usare come definizione provvisoria per $\mathbb{R}$ la seguente: $\mathbb{R}$ *è l'insieme di tutte le stringhe decimali senza però i limiti imposti per i numeri razionali.* Questo nuovo insieme ha la proprietà di contenere il vecchio insieme $\mathbb{Q}$, e di ereditare le sue operazioni interne $(+, \times, \leq)$ per le quali valgono sempre le stesse regole e proprietà. Inoltre adesso, con l'introduzione dei numeri reali, sulla retta numerica non vi sono più "buchi".

3.6.1 Limitato, limitato superiormente, limitato inferiormente

Prendiamo un insieme $A \subset \mathbb{R}$. A si dice *limitato superiormente* se $\exists M \in \mathbb{R}, \forall a \in A : a \leq M$. Detto in italiano, significa che A è limitato superiormente se esiste un numero maggiore di ogni elemento di A, il che corrisponde proprio alla definizione intuitiva del termine.

Non sempre si riesce a trovare questo valore M. Ad esempio, se si prende $A = \mathbb{N}$, non è possibile trovare alcun M. Ovviamente, se riesco a trovare anche un solo M, tutti gli $n \in \mathbb{R}, n > M$ possiedono la sua stessa proprietà. L'insieme di tutti questi valori, ovvero di tutti i valori che sono maggiori di ogni elemento di A, viene chiamato *insieme dei maggioranti*.

Ovviamente la definizione di *limitato inferiormente* è perfettamente analoga e speculare, ovvero $\exists m \in \mathbb{R}, \forall a \in A : m \leq a$. In questo caso, se prendiamo $A = \mathbb{N}$, allora abbiamo che esso è limitato inferiormente (ad esempio, -1 è minore di ogni naturale).

Se un insieme A è limitato sia superiormente che inferiormente, si suole dire in modo più generale che è *limitato*.

3.6.2 Massimo e minimo

$M \in \mathbb{R}$ si dice il massimo dell'insieme $A \subset \mathbb{R}$ se $M \in A \wedge \forall a \in A : a \leq M$. Molto importante è il fatto che M *deve appartenere ad* A. Inoltre, si può facilmente dimostrare che M, se esiste, è unico.

Prendiamo infatti, sempre ragionando per assurdo, che esistano due massimi dell'insieme A, M e M': ovviamente deve essere che $M \neq M'$. Quindi, dato che M è un massimo, ed M' per definizione di massimo appartiene ad A, allora è vero che $M' \leq M$. Inoltre, dato che M' è un massimo ed M appartiene ad A per definizione di massimo, $M \leq M'$. Ma, per la proprietà antisimmetrica, l'unica possibile conseguenza delle due affermazioni è che $M = M'$. Ma avevamo posto per ipotesi che $M \neq M'$ - siamo quindi entrati in contraddizione. Ecco quindi dimostrato che esiste un unico massimo.

Definire e dimostrare le proprietà del minimo è ovviamente perfettamente speculare. Inoltre, se A possiede massimo, è anche limitato superiormente, perché ovviamente il valore cercato nella definizione di 'limitato superiormente' è proprio il massimo. Tuttavia il contrario non è per forza vero.

Esempio 3.6.1 *Si prenda l'insieme* $A = \{x \in \mathbb{R} \mid x < 1\}$ *e si determinino massimo e/o minimo.*

Questo insieme è superiormente limitato in quanto $\exists M \in \mathbb{R}, \forall a \in A : a \leq M$*, e questo* M *può essere ad esempio* 1*. Tuttavia* 1 *non è il massimo dell'insieme, in quanto è troppo grande per appartenervi. Inoltre, un qualunque numero inferiore ad* 1*, diciamo* $1 - \varepsilon$*, è invece troppo piccolo per poter essere maggiore di tutti gli elementi di* A*, in quanto tutti i numeri tra* $1 - \varepsilon$ *ed* 1 *(come ad esempio* $1 - \frac{\varepsilon}{2}$*) sono maggiori di esso.*
Infine esso non è sicuramente inferiormente limitato, quindi A *non possiede nè massimo nè minimo.*

Un'altro esempio già più complesso può essere questo:

Esempio 3.6.2 *Eseguire le stesse operazioni dell'esempio precedente, con l'insieme:*

$$A = \left\{ \frac{n-1}{n+1} \mid n \in \mathbb{N} \right\}$$

Inanzitutto esaminiamo a grandi linee l'insieme dei valori di questo insieme:

$$-1,\ 0,\ \frac{1}{3},\ \frac{2}{4},\ \frac{3}{5},\ \frac{4}{6},\ \dots$$

Inanzitutto possiamo verificare che sicuramente tutti gli elementi di A *sono inferiori di 1, infatti il numeratore della frazione è sempre e comunque inferiore del denominatore:* $n - 1 < n + 1$*. Quindi, il loro rapporto è sempre minore di* 1*. Quindi l'insieme è superiormente limitato. Ora,* 1 *sicuramente non è il massimo dell'insieme perché non appartiene ad esso (abbiamo appena mostrato che* $\forall a \in A, a < 1$*), ma come possiamo sapere che questo numero non esiste? Esaminiamo la sequenza dei numeri. È facile mostrare che i numeri sono in ordine crescente, basta far vedere che* $p_n < p_{n+1}$*, infatti:*

$$\begin{aligned} \frac{n-1}{n+1} &< \frac{(n+1)-1}{(n+1)+1} \\ \frac{n-1}{n+1} &< \frac{n}{n+2} \\ (n+2)(n-1) &< n(n+1) \\ n^2 + n - 2 &< n^2 + n \end{aligned}$$

E quest'ultima eguaglianza è ormai ovvia. Quindi, per assurdo, se trovassimo un M *che è il massimo dell'insieme, avrebbe il suo corrispondente indice n (*$M = p_n$*). Basterà ora trovare l'elemento* p_{n+1} *e potremmo ricavare un numero più grande del massimo* M*, cosa che non può essere.* M *quindi non esiste.*

Per quanto riguarda il minimo invece, ricaviamo che questo è -1*; infatti* $p_0 = -1$*, ed abbiamo* $p_0 < p_1 < p_2 < \dots$*, e quindi* -1 *è un valore minore di tutti gli altri, e quindi è il minimo.*

3.6.3 Estremo superiore ed inferiore

Per definire i concetti di estremo superiore ed inferiore richiamiamo e formalizziamo il concetto di *maggiorante*. x è un maggiorante dell'insieme A se e solo se $\forall a \in A : a \leq x$. Ovviamente anche il massimo di un insieme, se esiste, è maggiorante di quell'insieme. Indichiamo con M_A l'insieme dei maggioranti, ovvero $\mathsf{M}_A = \{x \in \mathbb{R} \mid \forall a \in A : a \leq x\}$.

A questo possiamo definire il concetto di estremo superiore: *l'estremo superiore è il minimo dei maggioranti*. Dato che un importante teorema afferma che l'insieme dei maggioranti di un insieme superiormente limitato possiede sempre minimo, allora ogni insieme superiormente limitato possiede estremo superiore. L'estremo superiore di A si è soliti indicarlo come $\sup A$. Se A non è superiormente limitato, si scrive $\sup A = +\infty$. Ovviamente se $\sup A \in A \Rightarrow \sup A = \max A$.

Esempio 3.6.3 *Si cerchi l'estremo superiore dell'insieme A così definito:*

$$A = \{\forall x \in \mathbb{R} \mid x < 1\}$$

Qui $\mathsf{M}_A = \{\forall x \in \mathbb{R} \mid 1 \leq x\}$. *Infine, è facile mostrare che* $\min \mathsf{M}_A = 1$. *Quindi,* $\sup A = 1$.

Quale procedimento si può stabilire per decidere se un numero è o meno l'estremo superiore di un insieme A? Ovvero, come si risolve un problema in questa forma: *Dato $A \subset \mathbb{R}$ e $\lambda \in \mathbb{R}$, dimostrare che $\lambda = sup A$.*

Rispondere a questa domanda significa dimostrare due punti:

1. λ è un maggiorante di A ($\lambda \in \mathsf{M}_A$, ovvero $\forall x \in A : x \leq \lambda$).
2. λ è il minimo dell'insieme dei maggioranti, il che equivale anche a dire che è il suo estremo inferiore, dato che, come detto sopra, M_A possiede sempre minimo ($\lambda = \min \mathsf{M}_A$, ovvero $\lambda \in \mathsf{M}_A \ \wedge \ \forall m \in \mathsf{M}_A : \lambda \leq m$).

Per dimostrare questa seconda affermazione, si può procedere così. Se λ è il minimo di M_A, è come dire che un qualunque $\lambda - \varepsilon$, con $\varepsilon \in \mathbb{R}$, $\varepsilon > 0$, non è un maggiorante. Questo significa che la proposizione $\forall x \in A : x \leq \lambda - \varepsilon$ è falsa. In pratica, si deve dimostrare che è sempre possibile trovare un $x \in A$ più grande di $\lambda - \varepsilon$: $\exists x \in A : x > \lambda - \varepsilon$.
Facciamo un qualche esempio.

Esempio 3.6.4 *Si prenda nuovamente l'insieme A così definito:*

$$A = \left\{ \frac{n-1}{n+1} \mid n \in \mathbb{N} \right\}$$

E si dimostri che $\sup A = 1$.

In questo caso, quindi, $\lambda = 1$.

1. $\forall x \in A : x \leq 1$.
   Come già mostrato precendentemente, dato che il numeratore è sempre minore del denominatore, si ha che il loro rapporto è sempre minore di 1. *Quindi,* 1 *è un maggiorante di A.*
2. $\exists x \in A : x > 1 - \varepsilon$.

Noi sappiamo che possiamo esprimere ogni x di A come $(n-1)/(n+1)$: sostituiamo quindi la relazione all'interno dell'espressione da verificare e semplifichiamo:

$$\begin{aligned} \frac{n-1}{n+1} &> 1 - \varepsilon \\ \varepsilon &> 1 - \frac{n-1}{n+1} \\ \varepsilon &> \frac{n+1-n+1}{n+1} \\ \varepsilon &> \frac{2}{n+1} \\ n &> \frac{2}{\varepsilon} - 1 \end{aligned}$$

Cosa abbiamo fatto? In pratica abbiamo ricavato che valore deve avere n per soddisfare la relazione imposta. Ora, questa relazione è che n deve essere maggiore di un certo valore dipendente da ε: sicuramente, essendo che $n \in \mathbb{N}$, è sempre possibile trovare questo valore. È quindi anche sempre possibile trovare un corrispondente x che soddisfi la relazione. Quindi anche il secondo punto è stato dimostrato: 1 *è effettivamente l'estremo superiore dell'insieme.*

La definizione di estremo inferiore è analoga a quello di estremo superiore ma speculare:

$$\inf A \stackrel{\text{def}}{=} \max \mathsf{m}_A$$

m_A è l'insieme dei *minoranti* di A. La dicitura $\inf A = -\infty$ significa che A non è inferiormente limitato.
Stesso discorso vale per quanto riguarda l'esecuzione di un esercizio sull'estremo inferiore:

1. $\forall x \in A : \lambda \leq x$
2. $\forall x \in \mathsf{m}_A : x \leq \lambda$.
   Questa proposizione poi si può sempre trasformare in:
   $\exists x \in A : x < \lambda + \varepsilon$

Prendiamo ora un esercizio leggermente diverso:

Esempio 3.6.5 *Dato* $A = \{1, 1+\frac{1}{2}, 1+\frac{1}{2}+\frac{1}{3}, 1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}, \dots\}$ *dimostrare che* $\sup A = +\infty$.

Ovviamente il modo per mostrare questo, è di dimostrare il fatto che A non possiede maggioranti. Espresso simbolicamente:

$$\forall M \in \mathbb{R}, \exists x \in A : x > M$$

ovvero

$$\exists n \in \mathbb{N}, n \geq 1 \mid 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} > M$$

Purtroppo una disequazione del genere è irrisolvibile se espressa in questa forma. Questo mostra come non sempre (anzi, quasi mai) la risoluzione di un problema sugli estremi di un insieme sia facilmente risolvibile.
Ad esempio, per poter risolvere questo esempio bisogna ricorrere ad alcuni trucchi matematici. Osserviamo inanzitutto che, prendendo i primi 2 *elementi, si ha:*

$$1 + \frac{1}{2}$$

Prendendo i primi 4 *elementi:*

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4}$$

Tuttavia, dato che $\frac{1}{3} > \frac{1}{4}$*, abbiamo anche che* $\frac{1}{4} + \frac{1}{4} < \frac{1}{3} + \frac{1}{4}$*. Possiamo quindi dire che:*

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} > 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{4} = 1 + \frac{1}{2} + 2\frac{1}{4} = 1 + 2\frac{1}{2}$$

Prendiamo ora 8 *elementi, e con ragionamenti analoghi arriviamo a questo:*

$$\underbrace{1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4}}_{> 1 + 2\frac{1}{2}} + \underbrace{\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}}_{4\frac{1}{8}} = 1 + 3\tfrac{1}{2}$$

Quindi concludiamo che il numero formato da 8 *elementi è maggiore di* $1 + 3\frac{1}{2}$*. Proseguendo col ragionamento con* 16*,* 32*,* 64*, ecc ... elementi, ci si può convincere velocemente che con un numero formato da* 2^n *elementi è maggiore di* $1 + n\frac{1}{2}$*. Quindi, per tornare alla definizione iniziale, qualunque* M *io prenda, ci sarà sempre un numero nella forma* $1 + n\frac{1}{2}$ *più grandi di esso (basta prendere un* $n > 2(M-1)$*).Una volta trovato questo, ad esso corrisponderà un elemento di A che è a sua volta maggiore di esso, ovvero l'elemento di posizione* 2^n *- chiamiamolo* a_{2^n}*. Quindi, dato che* $M < 1 + n\frac{1}{2}$*, con* $n > 2(M-1)$*, e* $1 + n\frac{1}{2} < a_{2^n}$*, abbiamo che* $M < a_{2^n}$*. CVD.*

3.7 Funzione esponenziale in base a

La definizione di funzione esponenziale quando prendiamo $a > 0, a \in \mathbb{R}, p \in \mathbb{N}$ è piuttosto semplice:

$$a^0 = 1$$
$$a^p = \underbrace{a \times a \times a \times \dots}_{p \text{ volte}}$$

ovvero, in una forma ricorsiva più formale:

$$a^p = \begin{cases} 1 & \text{se } p = 0 \\ a \cdot a^{p-1} & \text{se } p > 0 \end{cases}$$

o ancora:

$$a^p = \prod_{i=1}^{p} a$$

Il nostro scopo, ora che abbiamo definito la funzione esponenziale per esponenti $\in \mathbb{N}$, è quello di estendere progressivamente l'appartenenza di p ad un insieme numerico.
Intanto, per definire $p \in \mathbb{Z}$, il compito è semplice, infatti sappiamo che sussiste la semplice uguaglianza:

$$a^{-p} = \frac{1}{a^p}$$

Che ci consente di calcolare anche potenze negative. Per estendere a $p \in \mathbb{Q}$ il compito è invece più complesso. Per farlo intanto bisogna introdurre questo teorema:

Teorema 3.7.1 *Dato* $n \geq 1,\ \forall y \in \mathbb{R},\ y \geq 0 :\ \exists!\ x \geq 0\ :\ x^n = y$

Ovvero, detto in termini più semplici, esiste una ed una sola radice ennesima positiva di un valore positivo o nullo. La radice n-esima di y, x, si può indicare come:

$$x = \sqrt[n]{y}$$

oppure

$$x = y^{\frac{1}{n}}$$

La dimostrazione risulta essere abbastanza complessa, e per questo dimostreremo prima quattro lemmi. In questi prenderemo sempre per ipotesi le stesse condizioni su n e che $x, y \geq 0$.

Lemma 3.7.2 $x^n \leq y^n \ \Leftrightarrow\ x \leq y$

Partiamo dall'esaminare la scomposizione di una differenza di potenze, prendendo per x ed y le stesse variabili considerate nel lemma:

$$\underbrace{x^n - y^n}_{a} = \underbrace{(x-y)}_{b}\underbrace{\left(x^{n-1} + x^{n-2}y + \cdots + xy^{n-2} + y^{n-1}\right)}_{c}$$

Dato che $x, y > 0$ ed $n > 1$, abbiamo allora che tutte le potenze del membro c sono positive, e la loro somma o prodotto è quindi positiva. Quindi, $c > 0$.
Inoltre, per ipotesi, $x^n \leq y^n$, quindi $x^n - y^n \leq 0$. $a \leq 0$.
Ora, se $a < 0$ allora dobbiamo ammettere che $b < 0$, e se $a = 0$ allora $b = 0$. Quindi $b \leq 0$, e quindi $x \leq y$. CVD.

Lemma 3.7.3 $x^n = y^n \ \Leftrightarrow\ x = y$

Dire che $x^n = y^n$ è equivalente a:

$$\begin{cases} x^n \leq y^n \\ y^n \leq x^n \end{cases}$$

Che, per il lemma 3.7.2, è come dire

$$\begin{cases} x \leq y \\ y \leq x \end{cases} \quad \Rightarrow \quad x = y$$

CVD.

Lemma 3.7.4 $x^n < y \ \Leftrightarrow\ \exists \varepsilon > 0\ :\ (x+\varepsilon)^n < y$

Questa parte è un po' più lunga delle altre. Il suo significato è che, se la potenza di un numero x è minore di un altro y, allora esiste sempre un altro numero $x + \varepsilon$, leggermente superiore ad x, la cui potenza rimane tuttavia minore di y.

Inanzitutto prendiamo in considerazione nella disequazione $(x+\varepsilon)^n < y$ solo il primo membro, e lo riscriviamo come:

$$(x+\varepsilon)^n \quad = \quad \underbrace{[(x+\varepsilon)^n - x^n]}_{(*)} \quad + \quad x^n$$

Il membro (*) può poi essere riscritto come

$$(\cancel{x} + \varepsilon - \cancel{x})[(x+\varepsilon)^{n-1} + (x+\varepsilon)^{n-2}x + \cdots + (x+\varepsilon)x^{n-2} + x^{n-1}]$$

Se riesco a dimostrare la disuguaglianza per $0 < \varepsilon < 1$, allora risulterà vera, dato che vale almeno per un ε. Quindi mi restringo volontariamente a questo intervallo. Possiamo quindi dire che:

$$x + \varepsilon < x + 1$$

e che

$$x < x + 1$$

Maggioro quindi (*), prima considerando che $x+\varepsilon < x+1$, poi che $x < x+1$:

$$\begin{aligned}(*) &< \varepsilon[(x+1)^{n-1} + (x+1)^{n-2}x + \ldots + (x+1)x^{n-2} + x^{n-1}] < \\ &< \varepsilon[(x+1)^{n-1} + (x+1)^{n-2}(x+1) + \ldots + (x+1)^{n-2}(x+1) + (x+1)^{n-1}] = \\ &= \varepsilon[(x+1)^{n-1} + (x+1)^{n-1} + \ldots + (x+1)^{n-1} + (x+1)^{n-1}] = \\ &= n\varepsilon(x+1)^{n-1}\end{aligned}$$

Ora, se riusciamo a dimostrare questa disuguaglianza:

$$(x+\varepsilon)^n \leq n\varepsilon(x+1)^{n-1} + x^n \leq y$$

(di cui abbiamo già dimostrato il primo membro) per almeno una ε, avremo anche dimostrato che $(x+\varepsilon)^n \leq y$. Ora, ricaviamo ε:

$$\varepsilon \leq \frac{y - x^n}{n(x+1)^{n-1}}$$

Disuguaglianza che è vera sempre per almeno una ε se il secondo membro è positivo. Ora, $n > 0$ per ipotesi, $(x+1)^{n-1}$ è potenza con base ed esponente positivo, quindi positivo. Infine, $y - x^n > 0$, perché per ipotesi $x^n < y$. Quindi, per almeno un ε è vera la disuguaglianza. CVD.

Lemma 3.7.5 $x^n > y \Leftrightarrow \exists \varepsilon > 0 \ : \ (x+\varepsilon) > 0, (x+\varepsilon)^n < y$

La dimostrazione di questo lemma è sulla falsariga del lemma 3.7.4.

Passiamo quindi ora al teorema vero e proprio.
Costruiamo quindi l'insieme A così definito:

$$A = \{a \in \mathbb{R} \mid a \geq 0, a^n < y\}$$

Il nostro scopo sarà ora dimostrare che la x cercata è in realtà $\sup A$.
Partiamo col mostrare che $0 \in A$ (ne è inoltre il minimo), quindi $A \neq \emptyset$, ed x, essendo stato preso come limite superiore e quindi come maggiorante, è > 0, che è un elemento di A. Inanzitutto bisogna quindi mostrare che A è superiormente limitato, e lo facciamo attraverso questa disequazione:

$$\forall a \in A : a^n < y < y+1 \leq (y+1)^n$$

Prendendo il primo e l'ultimo membro e applicando il lemma 3.7.2 si ricava che $a < y+1$; $y+1$ risulta quindi un maggiorante - l'insieme è superiormente limitato.

Per dimostrare che $x = \sup A$, si utilizzerà una dimostrazione per assurdo, ovvero si mostra che dalle supposizioni che $x^n < y$ oppure $x^n > y$ arrivo a delle contraddizioni, e che quindi l'unica realtà accettabile sia che $x^n = y$.

Parto supponendo che $x^n < y$: per il lemma 3.7.4 posso dire che

$$\exists \varepsilon > 0 \ : \ (x+\varepsilon)^n < y$$

$x+\varepsilon$ appartiene ad A per definizione, ma è anche vero che $x < x+\varepsilon$, fatto impossibile perché x è un maggiorante. Quindi questa prima ipotesi è inaccettabile.

Se suppongo che $x^n > y$, per il lemma 3.7.5 posso dire che

$$\exists \varepsilon > 0 \ : \ (x-\varepsilon) > 0, (x-\varepsilon)^n > y$$

D'altronde $y > a^n$ per la definizione di A, e quindi per il lemma 3.7.2 abbiamo che $\forall a \in A : x - \varepsilon > a$. $x - \varepsilon$ risulta quindi un maggiorante, ma $x > x - \varepsilon$, e da qui risulterebbe che x non è il minore dei maggioranti, il che è assurdo dato che che $x = \sup A$. Quindi anche questa seconda ipotesi va scartata.

Rimane quindi solo da supporre che $x^n = y$. E dato che l'estremo superiore di un insieme superiormente limitato come quello preso in considerazione esiste sempre ed è unico, abbiamo dimostrato il teorema.

C.V.D.

Già dalla stessa scrittura che abbiamo usato per indicare la radice n-esima, è facile dedurre il modo con cui definiamo la potenza ad esponente $\in \mathbb{Q}$:

$$a^{\frac{m}{n}} \ = \ \left(a^{\frac{1}{n}}\right)^m \ = \ \left(\sqrt[n]{a}\right)^m$$

Ora andiamo incontro all'ultimo ostacolo, che è estendere la definizione di a^p utilizzando $p \in \mathbb{R}$. Ad esempio, come definire $10^{\sqrt{2}}$? Possiamo arrivare a questa soluzione procedendo per approsimazioni successive. Noi sappiamo che possiamo rappresentare p come

$$p \ = \ \alpha_0, \alpha_1\alpha_2\alpha_3\ldots$$

con $\alpha_0 \in \mathbb{Z}$ e $\alpha_n \ (n > 0) \ \in \mathbb{N}$. Se prendiamo la successione dei valori *razionali* che si ottengono troncando p ad α_1 escluso, poi ad α_2, poi α_3 e così via, otteniamo un insieme i cui valori crescono sempre. È facile dimostrare che il valore p è l'estremo superiore di questo insieme.

Prendiamo quindi l'insieme delle potenze ottenute utilizzando come esponenti i valori di questo insieme, in sequenza:

$$P = \{\ a^{\alpha_0},\ a^{\alpha_0,\alpha_1},\ a^{\alpha_0,\alpha_1\alpha_2},\ \dots\ \}$$

Come vedremo, per una proprietà che diremo in seguito, anche questa successione è strettamente crescente, se prendiamo $a > 1$, o strettamente decrescente se $0 < a < 1$. Se $\alpha_0 < 0$, il discorso si inverte ma non cambia nella sua sostanza. Prendiamo il caso più semplice per il momento ($a > 1 \ \wedge \ \alpha_0 > 0$) e troviamo quindi che il valore a^p è il limite superiore dell'insieme così definito. Possiamo quindi dire:

$$a^p = \sup P$$

Abbiamo quindi costruito una funzione:

$$x \to a^x$$

Il cui dominio $D = \mathbb{R}$ e il codominio $C = \{x \in \mathbb{R} \mid x > 0\}$. Le sue principali proprietà (da cui è possibile ricavare tutte le altre) sono:

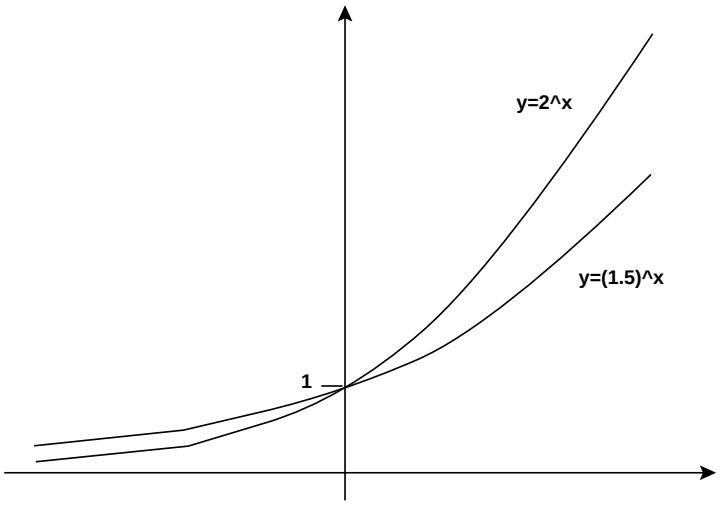
1. $\forall x \in \mathbb{R} : a^x > 0$
2. $\forall x, y \in \mathbb{R} : a^{x+y} = a^x a^y$
3. $\forall x, y \in \mathbb{R} : (a^x)^y = a^{xy}$
4. Per $a > 1$, cresce strettamente $(x' < x'' \Rightarrow a^{x'} < a^{x''})$
   Per $0 < a < 1$, descresce strettamente $(x' < x'' \Rightarrow a^{x'} > a^{x''})$
5. $\forall a, y > 0, \exists!\ x : a^x = y$

Considerando assieme la prima e la quinta proprietà, vediamo che ci dicono:

1. Tutti gli a^x sono positivi
5. Ogni numero positivo è un a^x

Esaminando la quarta proprietà inoltre si ricava che ogni a^x è una funzione *iniettiva*, ovvero $p, q \in D, p \neq q \Rightarrow f(p) \neq f(q)$.

Figura 3.3: Grafici delle esponenziali



3.8 Funzione logaritmica in base a

In questo paragrafo studieremo la funzione inversa dell'esponenziale - ovvero quella funzione che, dato un numero, resistuisce l'esponente a cui andrebbe elevata la base data per ottenere tale numero. Questa funzione viene chiamata *logaritmo in base a* e si indica con:

$$\lg_a y = x$$

(o anche $\log_a y = x$, noi utilizzeremo prevalentemente la prima notazione) dove abbiamo utilizzato gli stessi simboli dell'espressione $a^x = y$. Per avere una funzione, dobbiamo quindi studiare le soluzioni dell'equazione

$$a^x = y \quad a > 0$$

per quanto riguarda la x.

Se $a = 1$, allora è soddisfatta solo per $y = 1$, e qualunque x è la soluzione - se invece $y \neq 1$, nessun x risulta soluzione. In questo caso il logaritmo non potrebbe essere considerato una funzione (poichè a quanto risulta per $a = 1$, a^x non è una funzione iniettiva, che è condizione necessaria e sufficiente per costruirvi una funzione invertibile), e quindi poniamo $a \neq 1$. Vale comunque il seguente teorema:

Teorema 3.8.1 *Sia $a > 0$, $a \neq 1$, $y > 0$. Esiste allora un unico numero reale x tale che $a^x = y$, ovvero:*

$$\forall a, y \in \mathbb{R}; a > 0, a \neq 1, y > 0 \Rightarrow \exists! x : a^x = y$$

Le proprietà della funzione logaritmica possono essere dedotte da quelle della esponenziale e sono le seguenti:

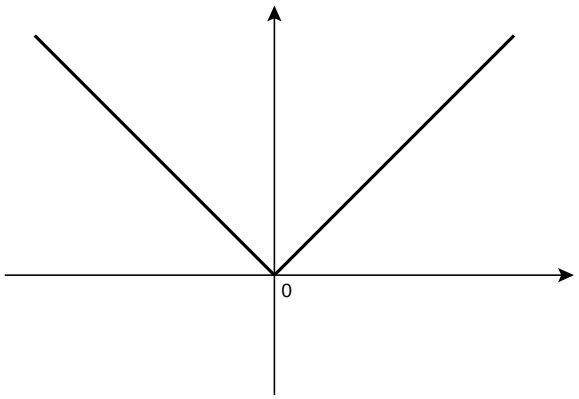
1. $\lg_a xy = \lg_a x + \lg_a y$
2. $\lg_a \dfrac{x}{y} = \lg_a x - \lg_a y$
3. $\lg_a x^\alpha = \alpha \lg_a x$
4. $\lg_a x = \dfrac{1}{\lg_x a} = -\lg_{\frac{1}{a}} x \quad (x \neq 1)$
5. $\lg_b x = \dfrac{\lg_a x}{\lg_a b}, \quad \forall x > 0, b \neq 1$

3.9 Valore assoluto e distanza di due numeri

Definiamo il valore assoluto $|x|$ (figura 3.4) come:

$$|x| = \begin{cases} x \text{ se } x \geq 0 \\ -x \text{ se } x < 0 \end{cases}$$

Figura 3.4: Grafico del valore assoluto



Definiamo inoltre la distanza tra due numeri A e B come:

$$dist_{A,B} \stackrel{\text{def}}{=} |A - B|$$

L'insieme, ad esempio, dei punti distanti meno di p dal punto a sono quindi definiti come:

$$\{x \mid |x - a| < p\}$$

Che, scritto in altro modo, è $(a - p, a + p)$. Se volessi invece avere anche gli estremi inclusi, utilizzerei la notazione $[a - p, a + p]$. Le proprietà di cui gode la distanza sono:

1. $dist_{A,B} \geq 0$
2. $dist_{A,B} = 0 \Rightarrow A = B$
3. $dist_{A,B} = dist_{B,A}$
4. $dist_{A,C} \leq dist_{A,B} + dist_{B,C}$ (*disuguaglianza triangolare*)

Inoltre possiamo scrivere che:

$$|x| = |x - 0| = dist_{x,0}$$

Ecco quindi che possiamo definire il valore assoluto in termini di distanza.

La disuguaglianza triangolare è formalmente equivalente a:

$$|x + y| \leq |x| + |y|$$

Infatti basta sostituire nell'espressione i valore $x = a - c$ ed $y = c - b$ per ottenere di nuovo la forma in cui è stata qui presentata. La dimostrazione della disuguaglianza triangolare in questa forma è abbastanza facile:

$$-|x| \leq x \leq |x| \quad \text{e} \quad -|y| \leq y \leq |y|$$

sommando membro a membro:

$$-(|x| + |y|) \leq x + y \leq |x| + |y|$$
$$|x + y| \leq |x| + |y|$$

come volevasi dimostrare.

3.10 Numeri Complessi

Sappiamo che per alcune equazioni di secondo grado, come per esempio $x^2 + 1 = 0$ non esiste alcuna soluzione, ovvero:

$$\nexists x \in \mathbb{R} \mid x^2 + 1 = 0$$

E ciò, in questo caso, è di facile dimostrazione:

$$x^2 \geq 0\ ,\ x^2 + 1 \geq 1\ ,\ x^2 + 1 \not< 1\ ,\ x^2 + 1 \neq 0$$

Tuttavia, per quanto riguarda altre equazioni, si hanno due soluzioni, come ad esempio $x^2 - 1 = 0$, ovvero:

$$\exists x_1, x_2 \in \mathbb{R} \mid x_1^2 - 1 = 0\ , x_2^2 - 1 = 0$$

E come ben sappiamo, $x_1 = 1,\ x_2 = -1$, oppure le due soluzioni invertite.
Questa disparità di risultati appare in un qualche modo una irregolarità dell'algebra, ma si può risolvere estendendo il campo in cui cercare le soluzione ad un insieme più vasto dei reali, chiamato insieme dei *numeri complessi* ($\mathbb{C}$).

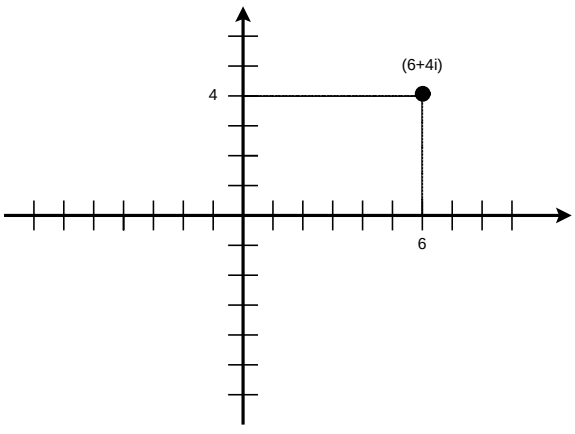
Indichiamo con $z = (a + ib)$ un qualunque numero di $\mathbb{C}$ - sia a che b appartengono a $\mathbb{R}$. a è chiamato parte *reale* di z, e b parte *immaginaria*. Quindi, formalmente, l'insieme dei numeri complessi non è altro che l'insieme di coppie di numeri reali: $\mathbb{C} \sim \mathbb{R} \times \mathbb{R}$.

Di solito si stabilisce una corrispondenza tra numeri reali e complessi, ovvero:

$$a \in \mathbb{R} \ \longleftrightarrow\ a + i0 \in \mathbb{C}$$

Per rappresentare i numeri reali ci siamo serviti di una retta, dato che essi sono definiti da un solo valore - in questo caso, essendo i numeri complessi definiti da due valori, ci serviremo di uno spazio cartesiano (figura 3.5).

Figura 3.5: Un numero complesso nel piano cartesiano



Definiamo ora le operazioni già definite in campo reale precedentemente: $+, \cdot, \leq$ e le loro relazioni.

$$(a + ib) + (\alpha + i\beta) = (a + \alpha) + i(b + \beta)$$

Si può facilmente verificare che la somma così definita è un'estensione di quella definita in campo reale:

$$a + b \ \rightarrow \ (a + i0) + (b + i0) = (a + b) + i0 \ \rightarrow \ a + b$$

Come si può vedere abbiamo compiuto la somma in campo complesso ottenendo lo stesso risultato che già avevamo in campo reale. Inoltre la somma così definita risulta possedere tutte le proprietà tipiche della somma: la proprietà associativa, commutativa e possiede elemento neutro $(0 + 0i)$, inoltre ogni elemento di $\mathbb{C}$ possiede opposto $(\forall z \in \mathbb{C}, \exists u \in \mathbb{C} \ : \ z + u = (0 + i0))$. In pratica, l'opposto di $(a + ib)$ è $(-a + i(-b))$.

$u \cdot v$ è già qualcosa di un po' più complicato. Proviamo ad eseguirla secondo le normali regole della moltiplicazione algebrica:

$$\begin{array}{l}(a + ib) \cdot (\alpha + i\beta) = \\ a\alpha + a \cdot i\beta + ib \cdot \alpha + ib \cdot i\beta = \\ a\alpha + i(a\beta + b\alpha) + i^2 \cdot b\beta\end{array}$$

Ed ora sorge il problema di interpretare il i^2. E qui sta il nucleo di ciò che sono i numeri complessi: infatti troviamo che la moltiplicazione risulta correttamente definita (come vedremo più avanti) se e solo se poniamo $i^2 = -1$, il che apparentemente può sembrare assurdo. Ma finora non abbiamo mai assunto che i sia un numero reale, ma semplicemente un 'segnaposto' per un qualche numero che segua le stesse proprietà dei reali per quanto riguarda le operazioni, come in effetti abbiamo verificato. Ora abbiamo che:

$$(a + ib) \cdot (\alpha + i\beta) = (a\alpha - b\beta) + i(a\beta + b\alpha)$$

Ed è facile verificare che anche la moltiplicazione, così definita, possiede tutte le proprietà che ci aspettiamo: è associativa, commutativa, possiede elemento neutro $(1 + i0$, il corrispondente di 1 in campo reale d'altronde) ed ogni numero complesso (escluso $0 + i0$) ha opposto. Vediamo di mostrare questi ultimi due punti.

Il problema di trovare l'elemento neutro in pratica si riassume nel risolvere per x e y la seguente equazione:

$$(a + ib)(x + iy) = (a + ib)$$

Ovvero

$$(ax - by) + i(ay + bx) = a + ib$$

Ora, se i due membri dell'equazione sono uguali, significa che hanno identiche parti reali e immaginarie:

$$\left\{ \begin{array}{l} ax - by = a \\ ay + bx = b \end{array} \right. \qquad \left\{ \begin{array}{l} a(x - 1) = by \\ b(x - 1) = -ay \end{array} \right.$$

Senza entrare nella risoluzione di un sistema lineare nei dettagli, il risultato si può trovare velocemente. Se quest'ultima equazione è vera per ogni $(a + ib)$, deve essere vera anche in particolare per $a = 0, b = 1$, da cui ricavo $y = 0$, e per $a = 1, b = 0$, da cui ricavo $x = 1$. Sostituendo questi valori nel sistema ottengo un sistema sempre vero, e quindi la soluzione trovata è effettivamente giusta. Inoltre l'elemento neutro, se esiste, è unico, quindi ho anche trovato *tutte* le soluzioni. Ecco quindi che ho dimostrato che l'elemento neutro per la moltiplicazione è $(1 + i0) \in \mathbb{C} \leftrightarrow 1 \in \mathbb{R}$. Troviamo ora l'opposto di un generico $(a + ib) \neq (0 + i0)$.

$$(a + ib)(x + iy) = 1 + i0 \quad (a + ib) \neq (0 + i0)$$

$$\left\{ \begin{array}{l} ax - by = 1 \\ ay + bx = 0 \end{array} \right.$$

$$\left\{ \begin{array}{l} a^2x - aby = a \\ aby + b^2x = 0 \end{array} \right. \qquad \textit{moltiplico entrambi i membri per a}$$

$$\left\{ \begin{array}{l} // \\ (a^2 + b^2)x = a \end{array} \right. \qquad \textit{sommo membro a membro}$$

$$x = \frac{a}{a^2 + b^2}$$

Oppure, se invece di moltiplicare per a, moltiplico per $-b$:

$$\left\{ \begin{array}{l} abx - b^2y = b \\ -a^2y - abx = 0 \end{array} \right. \qquad \textit{moltiplico entrambi i membri per -b}$$

$$\left\{ \begin{array}{l} // \\ -(a^2 + b^2)y = b \end{array} \right. \qquad \textit{sommo membro a membro}$$

$$y = -\frac{b}{a^2 + b^2}$$

Quindi:

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

Anche questa operazione ha un corrispondente in campo reale, infatti se calcoliamo l'inverso di $(a+i0)$ risulta proprio uguale ad $1/a$.

Ora rimane tuttavia un'ultima caratteristica che in campo reale era presente e che non abbiamo ancora considerato, ovvero gli operatori d'ordine. Ci chiediamo quindi se esiste un ordine nel campo complesso esprimibile con l'operatore $\leq$:

$$a+ib \overset{?}{\leq} \alpha+i\beta$$

Ricordiamoci che questo operatore deve possedere le tre proprietà di riflessività, antisimmetria e transitività.

Come primo tentativo potremmo essere tentati a provare la semplice doppia relazione:

$$a+ib \leq \alpha+i\beta \overset{\text{def}}{=} \begin{cases} a \leq \alpha \\ b \leq \beta \end{cases}$$

Ma è facile notare che, sebbene la relazione così definita sia riflessiva e transitiva, non è antisimmetrica: un esempio che mostra ciò è il seguente:

$$a+ib = 1+3i, \quad \alpha+i\beta = -1+5i$$

e qui troviamo che:

$$1+3i \not\leq -1+5i \text{ perché } 1 \not\leq -1$$
$$-1+5i \not\leq 1+3i \text{ perché } 5 \not\leq 3$$
$$1+3i \neq -1+5i$$

Possiamo quindi provare il corrispondente di un ordine lessicografico: ovvero si confrontano prima le componenti reali dei numeri e, se sono uguali, in seguito quelle immaginarie.

Con una relazione così definita (indichiamola con $\preceq$) si trova che effettivamente è una relazione d'ordine, però le manca ancora qualcosa. Infatti in campo reale le operazioni + e $\times$ sono correlate alla relazione d'ordine $\leq$ da due proprietà:

$$a \leq b \Rightarrow a+c \leq b+c$$
$$a \leq b \Rightarrow ac \leq bc, c > 0$$

E queste dalla relazione $\preceq$ non sono soddisfatte. Quindi, neanche $\preceq$ è ciò che cercavamo.

È possibile dimostrare, in realtà, che *non esiste nessuna relazione d'ordine correlata con le operazioni in campo complesso.* Infatti, proprio partendo dalle proprietà che collegano tra loro relazioni d'ordine ($\leq$) e somma e moltiplicazione:

$$a \leq b \Rightarrow ac \leq bc, \text{se } c \geq 0$$

Si ricava, ponendo $a = m$, $b = m+m$ e $c = m$ ed $m \geq 0$ (per $m < 0$ il ragionamento è sostanzialmente uguale):

$$\begin{aligned} mm &\leq (m+m)m \\ mm &\leq mm+mm \\ (mm-mm) &\leq mm+(mm-mm) \\ 0 &\leq mm \end{aligned}$$

Il che in pratica ci dice "in un sistema in cui valga la proprietà sopracitata, il quadrato di un elemento non può mai essere minore di zero, ovvero negativo". Il che, come noi sappiamo, in campo complesso è falso (ad esempio i è un elemento di $\mathbb{C}$ il cui quadrato è *negativo*). Quindi, la proprietà da cui siamo partiti è, in campo complesso, falsa, e dunque dobbiamo accettare di perdere una relazione d'ordine con buone proprietà come quelle che avevamo trovato in $\mathbb{R}$.

Tuttavia, non tutto ciò che deriva dall'idea di ordine è perso. Ad esempio esiste la nozione di distanza, che è intesa proprio come la distanza tra due punti nel campo cartesiano complesso.

$$dist_{a+ib,\alpha+i\beta} \overset{def}{=} \sqrt{(a-\alpha)^2+(b-\beta)^2}$$

Inoltre, così come abbiamo in fatto in campo reale, possiamo scrivere questa uguaglianza:

$$dist_{A,B} = |A-B|$$
$$|a+ib| = dist_{a+ib,0} = \sqrt{a^2+b^2}$$

Abbiamo così definito il valore assoluto in campo complesso. Pur essendo i termini 'valore assoluto' e 'modulo' teoricamente sinonimi, si preferisce utilizzare il primo in campo reale e il secondo in campo complesso.

Le proprietà in campo complesso dei valori assoluti sono le stesse che già abbiamo trovato in campo reale:

1. $|z-u| \geq 0$

2. $|z-u|=0 \Rightarrow z=u$
3. $|z-u|=|u-z|$
4. $|z-u| \leq |z-w|+|w-u|$

Quest'ultima disuguaglianza, il cui nome è triangolare come sappiamo già da prima, ora assume anche un senso geometrico, se la si osserva nel piano cartesiano, ovvero dice che la lunghezza del lato di un triangolo non è mai maggiore della somma degli altri due. Ecco quindi anche spiegato il suo nome.

Per definire i numeri complessi in questa forma (detta *cartesiana*) si è in pratica partiti dall'assunto che $i^2=-1$ e da lì si sono definite le varie proprietà di questi numeri. Un altro modo per fare ciò può essere definire l'insieme dei numeri complessi come prodotto cartesiano dei reali con sè stessi ($\mathbb{R} \times \mathbb{R}$). In questo modo si hanno varie coppie ordinate che possono essere interpretate come le coppie parte reale - parte immaginaria dei numeri complessi. Se a questo punto si pongono *per definizione* (e non come conseguenza) le operazioni di somma e prodotto come quelle da noi riportate, si può arrivare quindi alla conclusione che $i^2=-1$:

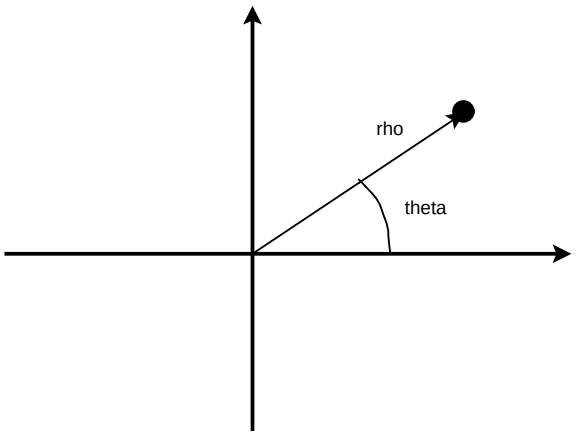
$$(a,b)\cdot(c,d) \stackrel{def}{=} (ac-bd, ad+bc)$$
$$(0,1)\cdot(0,1) = (0\cdot 0 - 1\cdot 1, 0\cdot 1 + 1\cdot 0) = (-1,0)$$

Quindi dalla definizione delle operazioni si arriva a definire i, mentre la strada che abbiamo noi seguito qui fa la strada inversa. Entrambe sono valide, e dato che si possono definire l'una nei termini dell'altra, hanno la stessa potenza espressiva e conducono agli stessi risultati. Sono presenti anche vari altri modi di definire i numeri complessi, che conducono tutti agli stessi risultati.

3.10.1 Rappresentazione polare dei complessi

Così come è possibile rappresentare le coordinate di un punto sia in forma cartesiana che *polare* (ovvero indicando la distanza dall'origine degli assi e l'angolo che forma la congiungente origine-punto con il semiasse positivo delle x), si può fare la stessa cosa con i numeri complessi (figura 3.6).

Figura 3.6: Rappresentazione polare di un numero complesso



Preso il numero complesso z, solitamente si indica con ρ il *modulo* di z, ovvero la distanza del punto che rappresenta z sull'asse cartesiano dall'origine degli assi, e con θ l'angolo, o *argomento*. Le relazioni per passare da un sistema all'altro sono facilmente ricavabili con qualche conoscenza di trigonometria e geometria cartesiana:

$$\begin{cases} \rho = dist_{z,0} = |z| = \sqrt{a^2+b^2} \\ \sin\theta = b/\rho \\ \cos\theta = a/\rho \end{cases}$$

Da cui si ricava

$$\begin{cases} a = \rho\cos\theta \\ b = \rho\sin\theta \end{cases}$$

Ovviamente, nella forma polare l'argomento è sempre determinato a meno di 2π – questa indeterminazioni di solito non porta problemi, ma nel caso in cui serva un unico valore si specifica l'intervallo in cui cercarlo, solitamente $[0, 2\pi)$ oppure $[-\pi, +\pi)$.

Esempio 3.10.1 *Convertire $z=1+i$ in forma polare.*

abbiamo inanzitutto che $\rho = \sqrt{a^2+b^2} = \sqrt{1^2+1^2} = \sqrt{2}$. inoltre:

$$\sin\theta = 1/\sqrt{2} = \sqrt{2}/2$$
$$\cos\theta = 1/\sqrt{2} = \sqrt{2}/2$$

Da cui ricavo $\theta = \pi/4$. Un modo per scrivere un numero in forma polare potrebbe essere questo:

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

Quindi, nel nostro caso: $1 + i = \sqrt{2}(\cos\pi/4 + i\sin\pi/4)$.

Inoltre (più avanti si scoprirà perché), il modo più usato per scrivere un numero complesso in forma polare è questo:

$$a + ib = \rho e^{i\theta}$$

C'è da sottolineare che questo non è solo un 'giochetto tipografico', ma ha un suo vero senso matematico. Espresso in questa forma, il nostro numero diventerebbe:

$$1 + i = \sqrt{2}\, e^{i\frac{\pi}{4}}$$

La rappresentazione polare porta alcuni vantaggi. Uno di questi, ad esempio, è quello di riuscire finalmente ad interpretare geometricamente il prodotto di due numeri complessi. Infatti, se li scriviamo in forma polare:

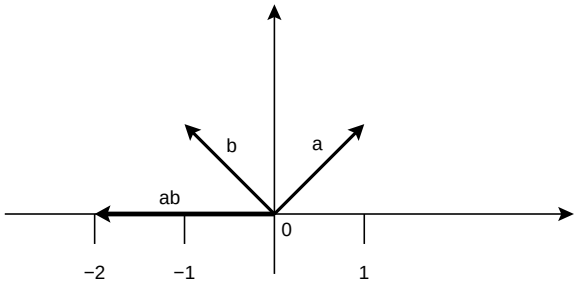
$$z = \rho(\cos\theta + i\sin\theta)$$
$$z' = \rho'(\cos\tau + i\sin\tau)$$

Abbiamo che:

$$z \cdot z' = \rho\rho'[(\cos\theta\cos\tau - \sin\theta\sin\tau) + i(\sin\theta\cos\tau + \cos\theta\sin\tau)] =$$
$$\rho\rho'[\cos(\theta + \tau) + i\sin(\theta + \tau)]$$

Vediamo quindi il risultato nel piano cartesiano in figura 3.7.

Figura 3.7: Prodotto tra numeri complessi



Con numeri complessi di modulo 1 si ha una vera e propria *rotazione* dell'uno dell'angolo dell'altro numero. Con numeri di modulo diverso da uno si ha invece anche una *dilatazione* di un fattore pari al modulo dell'altro numero.

Se quindi proviamo a calcolare la potenza di un numero complesso:

$$z^n = \underbrace{z \cdot z \cdot \ldots \cdot z}_{n \text{ volte}} = \rho^n(\cos(n\theta) + \sin(n\theta)) = \rho^n \cdot e^{in\theta}$$

Questa formula è spesso chiamata anche *formula di De Moivre.*

Passiamo ora a definire cos'è il *coniugato* di un numero complesso, indicato con $\overline{z}$.

$$z = a + ib$$
$$\overline{z} = a - ib$$

Geometricamente parlando, $\overline{z}$ è il simmetrico rispetto all'asse x del vettore z. Proviamo a calcolare ora z^{-1}:

$$z^{-1} = \frac{1}{z} = \frac{a}{a^2 + b^2} - i\frac{b}{a^2 + b^2} = \frac{a - ib}{a^2 + b^2} = \frac{\overline{z}}{|z|^2}$$

Oppure, espresso in forma polare:

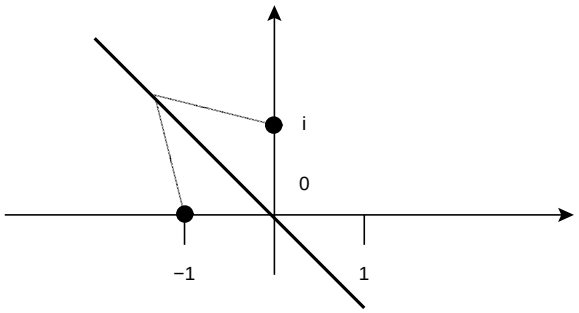
$$\overline{z} = \rho[\cos(-\theta) + i\sin(-i\theta)] = \rho(\cos\theta - i\sin\theta)$$

E quindi:

$$\frac{1}{z} = \frac{\rho(\cos\theta - i\sin\theta)}{\rho^2} = \frac{1}{\rho}(\cos\theta - i\sin\theta)$$

Si ha quindi che, nel caso di $\rho = 1$, il coniugato di z corrisponde al suo inverso, altrimenti si trova sempre sulla stessa direttrice, ma il suo modulo è il reciproco del modulo di z.

Figura 3.8: Luogo dei punti tali che $|z-i| = |z+1|$



Esempio 3.10.2 *Disegnare l'insieme dei z complessi tali che* $|z-i| = |z+1|$.

Geometricamente parlando, si tratta di trovare l'insieme dei punti equidistanti dai punti i e -1*, quindi di coordinate* $(0,1)$ *e* $(-1,0)$*. Questi punti sono ovviamente la bisettrice del secondo e quarto quadrante (figura 3.8).*

Verifichiamolo ora algebricamente:

$$\begin{aligned} |z-i| &= |z+1| \\ (z &= a+ib) \\ |a+i(b-1)| &= |(a+1)+ib| \\ \sqrt{a^2+b^2-2b+1} &= \sqrt{a^2+2a+1+b^2} \\ \cancel{a^2}+\cancel{b^2}-2b+\cancel{1} &= \cancel{a^2}+2a+\cancel{1}+\cancel{b^2} \\ -2a-2b &= 0 \\ a &= -b \end{aligned}$$

Quindi, espresso in forma cartesiana:

$$y = -x$$

oppure come insieme:

$$I = \{z = x - ix \mid x \in \mathbb{R}\}$$

Esempio 3.10.3 *Risolvere l'equazione* $z^4 = |z|$ *in campo reale e complesso, e confrontare i due risultati.*

Partiamo col risolvere in campo reale:

$$x^4 = |x|$$

Se $x = 0$*, l'equazione è verificata, se* $x \neq 0$*:*

$$x \neq 0 : \begin{cases} x > 0 : & x^4 = x \ \rightarrow \ x^3 = 1 \ \rightarrow \ x = 1 \\ x < 0 : & x^4 = -x \ \rightarrow \ x^3 = -1 \ \rightarrow \ x = -1 \end{cases}$$

Quindi troviamo le tre soluzioni: $-1, 0, 1$.

Se ora passiamo al caso complesso, possiamo inanzitutto considerare che le tre soluzioni reali sono senz'altro anche soluzioni complesse: $-1+i0, 0+i0, 1+i0$*. Per trovarle però tutte, ci conviene elaborare di nuovo le equazioni, utilizzando z scritto in forma polare (accorgimento spesso utile quando appaiono dei valori assoluti):*

$$\begin{aligned} z &= \rho(\cos\theta + i\sin\theta) \\ z^4 &= \rho^4[\cos(4\theta) + i\sin(4\theta)] \end{aligned}$$

L'equazione diventa quindi:

$$\begin{aligned} z^4 &= |z| \\ \rho^4[\cos(4\theta) + i\sin(4\theta)] &= \rho \end{aligned}$$

Ora, è facile capire che due numeri espressi in forma polare sono uguali se e solo se i loro moduli sono uguali e gli argomenti differiscono per multipli di 2π*. Quindi:*

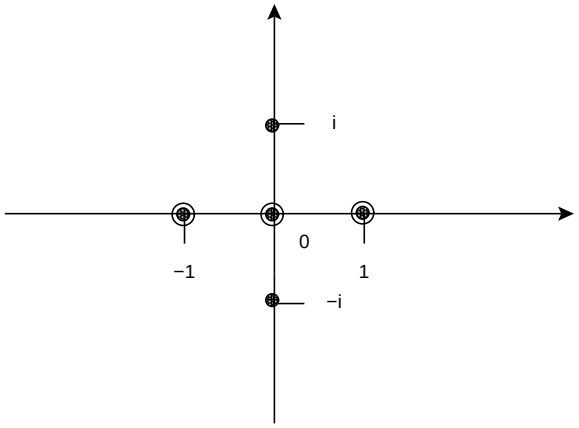
$$\begin{aligned} \rho^4 = \rho, \rho > 0 \quad e \quad & 4\theta = 0 + 2k\pi, k \in \mathbb{Z} \\ \rho = 0 \vee \rho = 1 \qquad & 2\theta = k\pi \\ & \theta = k\frac{\pi}{2} \end{aligned}$$

Apparentemente, sono quindi presenti infiniti numeri complessi che soddisfano questa equazione: in realtà, bisogna considerare la molteplicità dei risultati, ovvero più risultati che rappresentano lo stesso numero – nell'arco dei 2π troviamo infatti solo quattro angoli che soddisfano l'equazione:

$$\theta = \begin{cases} \pi/2 \\ \pi \\ 3\pi/2 \\ 2\pi \end{cases}$$

E quindi, 4 numeri complessi con modulo 1 ed angoli sopra indicati, ed un ulteriore numero con modulo 0. La loro rappresentazione cartesiana è quindi: $(1+i0), (0+i), (-1+i0), (0-i), (0+i0)$ *(figura 3.9).*

Figura 3.9: Esercizio



Come insieme:

$$I = \{-1, 0, 1, i, -i\}$$

Quindi, nel campo complesso abbiamo trovato due soluzioni in più rispetto al caso reale.

Esempio 3.10.4 *Analizzare la tipica equazione di secondo grado:*

$$ax^2 + bx + c = 0$$

In campo complesso, ovvero con $x \in \mathbb{C}$.

Utilizzando le stesse trasformazioni che si utilizzano in campo reale:

$$ax^2 + bx + c = a\left(x + \frac{b}{2a}\right)^2 + \left(c - \frac{b^2}{4a}\right) \longrightarrow \left(x + \frac{b}{2a}\right)^2 = \frac{b^2 - 4ac}{4a^2}$$

Ora, in campo reale ci si doveva fermare quando il secondo membro era un numero negativo, ovvero se il cosidetto determinante *($\Delta = b^2 - 4ac$) era negativo, perché in quel caso non esisteva nessun numero in campo reale il cui quadrato fosse un numero negativo. Ma in campo complesso abbiamo già visto come questo non sia vero (ad es., $i^2 = -1$).*

Quindi, in sostanza, per andare avanti il problema che ci troviamo a dover risolvere è:

$$z^2 = k, \quad z, k \in \mathbb{C}$$

Esprimiamo quindi z e k con la rappresentazione polare:

$$z = \rho(\cos\theta + i\sin\theta)$$
$$k = \rho'(\cos\tau + i\sin\tau)$$
$$z^2 = \rho^2[\cos(2\theta) + i\sin(2\theta)]$$

Come al solito, poniamo i moduli uguali e positivi e gli argomenti uguali a meno di multipli di 2π.

$$\rho^2 = \rho' \quad e \quad 2\theta = \tau + 2k\pi$$
$$\rho = \sqrt{\rho'} \qquad \theta = \frac{\tau}{2} + k\pi$$

Troviamo quindi in $[0, 2\pi)$ *due soluzioni:*

$$z = \begin{cases} \sqrt{\rho'}(\cos(\tau/2) + i\sin(\tau/2)) \\ \sqrt{\rho'}(\cos(-\tau/2) + i\sin(-\tau/2)) \end{cases}$$

ovvero:

$$z_1 = \pm\sqrt{\rho'}\left(\cos\frac{\tau}{2} + i\sin\frac{\tau}{2}\right)$$
$$z_2 = \overline{z_1}$$

Troviamo quindi che le due soluzioni sono coniugate (graficamente sono quindi simmetriche rispetto all'asse y). Il fatto più interessante è che queste due soluzioni esistono sempre*, qualunque sia* $k \in \mathbb{C}$*. Se ad esempio cerchiamo le soluzioni dell'equazione:*

$$z^2 = -1$$

Troviamo:

$$-1 \longrightarrow \rho = 1, \theta = \pi$$
$$z = \pm\sqrt{1} \cdot e^{\pi/2} = \pm i$$

Interessante il fatto che si è tornati alla definizione di i: ovvero, il numero il cui quadrato è pari a -1*. Più interessante è il fatto che non è l'*unico *numero il cui quadrato è* -1*: anche* $-i$ *possiede questa proprietà.*

A questo punto possiamo tornare al problema originale, ovvero risolvere l'equazione

$$\left(x + \frac{b}{2a}\right)^2 = \frac{\Delta}{4a^2}$$

in campo complesso. Ora, noi dobbiamo prendere il caso in cui $\Delta < 0$*, che era il caso che in campo reale era rimasto insoluto. In questo caso,* $x + b/2a$ *sarà uguale al numero complesso:*

$$\rho = \sqrt{-\Delta/(4a^2)}$$
$$\theta = 0/2 = 0$$

da cui ricavo come soluzione finale:

$$x = -\frac{b}{2a} \pm i\frac{\sqrt{-\Delta}}{2a}$$

Da notare il fatto che le due radice dell'equazioni sono numeri complessi coniugati. Si può inoltre dimostrare che questa equazione vale anche per $a, b, c \in \mathbb{C}$*. Infine, si può generalizzare questo risultato ad una qualunque equazione algebrica nella forma "polinomio uguale a zero", ovvero:*

Teorema 3.10.1 (Teorema fondamentale dell'algebra) *Ogni equazione composta da un polinomio di grado* $n \geq 1, n \in \mathbb{N}$ *eguagliato a 0, possiede n radici complesse.*

Esempio 3.10.5 *Risolvere l'equazione*

$$x^6 - 1 = 0$$

in campo complesso ($x \in \mathbb{C}$*).*

Inanzitutto, possiamo compiere questi semplici passaggi per trovare le soluzioni in campo reale:

$$x^6 = 1$$
$$x = \pm 1$$

In campo complesso invece conviene come sempre esprimere l'incognita in forma polare:

$$x = \rho(\cos\theta + i\sin\theta)$$
$$x^6 = \rho^6[\cos(6\theta) + i\sin(6\theta)]$$
$$1 = \rho^6[\cos(6\theta) + i\sin(6\theta)]$$

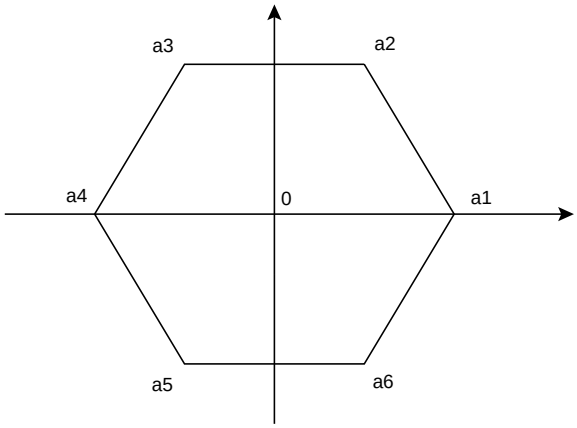
Come al solito confrontiamo moduli ed argomenti

$$\rho^6 = 1,\ \rho \geq 0 \quad 6\theta = 2k\pi,\ k \in \mathbb{Z}$$
$$\rho = 1 \quad \theta = k\frac{\pi}{3}$$

Compiendo le solite considerazione sui numeri complessi che vengono ripetuti più volte, troviamo:

$$z = \begin{cases} 1 \cdot (\cos 0 + \sin 0) = 1 + i0 \\ 1 \cdot (\cos \pi/3 + \sin \pi/3) = 1/2 + i\sqrt{3}/2 \\ 1 \cdot (\cos 2\pi/3 + \sin 2\pi/3) = -1/2 + i\sqrt{3}/2 \\ 1 \cdot (\cos \pi + \sin \pi) = -1 + i0 \\ 1 \cdot (\cos 4\pi/3 + \sin 4\pi/3 = -1/2 - i\sqrt{3}/2 \\ 1 \cdot (\cos 5\pi/3 + \sin 5\pi/3) = 1/2 - i\sqrt{3}/2 \end{cases}$$

Figura 3.10: Radici seste dell'unità



Quello che abbiamo fatto in pratica è stato trovare le radice n-esime (in questo caso, con $n = 6$) dell'unità in campo complesso. Quel che è risultato è che, in accordo con il teorema fondamentale dell'algebra (3.10.1), vi sono 6 radici. Inoltre, si può capire dal disegno, le radici n-esime dell'unità disegnano in campo complesso un n-agono (in questo caso, un esagono).

Abbiamo, nel corso di questi esempi, notato vari dei vantaggi che portano i numeri complessi, che si possono poi sostanzialmente ridurre al fatto che *in campo complesso è sempre possibile l'operazione di estrazione della radice n-esima.* Inoltre, si potrebbe vedere che anche l'operazione di logaritmo ha il dominio allargato a tutti i numeri complessi eccetto quelli corrispondenti alla semiretta negativa dei reali – comunque una notevole estensione del dominio!

Capitolo 4

Successioni e serie

4.1 Successioni

4.1.1 Definizione di successione

Una successione è una corrispondenza tra i numeri naturali (oppure un sottoinsieme dei numeri naturali presi da un certo n_0 in poi) e un insieme di numeri naturali. Quindi, una successione puó essere vista come una funzione da $\mathbb{N}$ a $\mathbb{R}$.

La legge della successione è quindi quella legge che indica come legare un numero naturale al suo corrispondente.

Alcuni esempi di successioni possono essere i seguenti:

$$\left(\tfrac{n-1}{n+1}\right)_{n\in\mathbb{N}}$$

$$\left(2^{1/n}\right)_{n\in\{x\in\mathbb{N}\mid x\geq 1\}}$$

$$\left((-1)^n\right)_{n\in\mathbb{N}}$$

Ad esempio nella seconda successione il dominio è un sottoinsieme di $\mathbb{N}$ che parte da $n_0 = 1$.

Si dice che una successione a_n è *limitata inferiormente* se $\exists m : \forall n : a_n > m$, ovvero se l'insieme dei valori di a_n è limitato inferiormente. Analogamente si definisce il concetto di *limitata superiormente*. Una successione limitata sia superiormente che inferiormente si dice genericamente *limitata*.

4.1.2 Limite di una successione per $n \to +\infty$

Una delle informazioni che sono utili sapere su una successione, è questa: come si comportano i valori successivi che assume a_n man mano che n aumenta? O, in altre parole, a che valore si avvicinano (se mai si avvicinano ad un valore) i successivi a_n? Tale valore, espresso per ora in maniera alquanto intuitiva, è detto *limite della successione per n che tende a* ∞ e si puó indicare con le seguenti notazioni:

$$\begin{aligned}\lim_{n\to\infty} a_n &= l\\ a_n &\overset{n\to\infty}{\longrightarrow} l\\ a_n &\longrightarrow l\end{aligned}$$

Talvolta ció puó essere intuitivo analizzando i successivi valori della successione, altre volte molto di meno. Quel che serve in realtá è trovare una definizione che ci consenta di esprimere il fatto che la successione *converga* ad un certo valore, quando si prendono termini sempre piú avanti.

In pratica quello che si vuole dire è che i vari a_n si avvicinano sempre di piú ad un certo valore l man mano che n aumenta. Questo può essere detto in questo modo: la distanza tra a_n ed l diventa sempre piú piccola (ovvero piccola a piacimento) quando n è abbastanza grande. O ancora, dato un qualunque "errore" da assegnare ad l, tutti i punti della successione a_n dopo un certo n_ε sono compresi al suo interno. Scritto sotto forma di formula:

$$\forall\varepsilon\in\mathbb{R}, \varepsilon > 0, \exists n_\varepsilon \in \mathbb{N} \ : \ |a_{n>n_\varepsilon} - l| < \varepsilon$$

E questa è la definizione di *limite finito di una successione per n che tende a* $+\infty$. L'espressione $|a_{n>n_\varepsilon} - l| < \varepsilon$ puó anche essere riformulata come:

$$a_n \in \,]l-\varepsilon, l+\varepsilon[, \ n > n_\varepsilon$$

Quel che rende complessa questa definizione e la sua verifica è l'arbitrarietá di ε.

É facile dimostrare che:

Teorema 4.1.1 (Teorema sull'unicità del limite) *Il limite di una successione, se esiste, è unico.*

Prendiamo per assurdo che esistano due limiti l_1 ed l_2 ($l_1 \neq l_2$) per la successione a_n. Allora sará possibile trovare almeno un ε per il quale i due intervalli $]l_1 - \varepsilon, l_1 + \varepsilon[$ e $]l_2 - \varepsilon, l_2 + \varepsilon[$ non si intersecano. Per questo ε sicuramente non sará possibile che la condizione di limite valga sia per l_1 che per l_2, in quanto ogni punto di a_n puó stare in un intervallo o nell'altro, ma mai in entrambi assieme. Quindi, l_1 non puó essere diverso da l_2, il che è come dire che il limite è unico. CVD.

Da questa osservazione si può poi anche dedurre un altro fatto, ovvero che *ogni successione convergente è limitata.*

Esempio 4.1.1 *Preso $0 < \alpha < 1$, $a_n = \alpha^n$, verificare che $a_n \to 0$.*

Per verificare ció dobbiamo mostrare che per ogni $\varepsilon > 0$, qualunque esso sia, esiste sempre un indice n_ε della successione a_n tale che, dopo di lui, $|a_n - l| < \varepsilon, n > n_\varepsilon$.

In questo caso abbiamo $l = 0$. Il membro sinistro della disuguaglianza diventa quindi $-\varepsilon < a_n < \varepsilon$, ovvero:

$$\begin{cases} -\varepsilon < \alpha^n \\ \alpha^n < \varepsilon \end{cases}$$

Ora, la prima uguaglianza è sempre verificata: $-\varepsilon$ è sempre negativo, α^n è invece sempre positivo. Sulla seconda possiamo invece compiere la semplice operazione di applicare ad entrambi i membri il logaritmo in base α, ricordando che a queste condizione esso è una funzione decrescente*, e quindi costringe ad invertire il verso della disuguaglianza:*

$$\begin{aligned} \ln_\alpha \alpha^n &> \ln_\alpha \varepsilon \\ n &> \ln_\alpha \varepsilon \end{aligned}$$

Dato che ε è positivo, esisterá sempre il logaritmo, e quindi un n cosí definito. Ecco che abbiamo trovato il nostro n_ε.

Esempio 4.1.2 *Dato*

$$a_n = \frac{n^2 + 2}{2n^2 + 1}$$

dimostrare che $a_n \to \dfrac{1}{2}$.

Come prima, impostiamo il sistema di disuguaglianze e sviluppiamolo

$$\begin{aligned} \frac{1}{2} - \varepsilon < \frac{n^2 + 2}{2n^2 + 1} &< \varepsilon + \frac{1}{2} \\ -\varepsilon < \frac{n^2 + 2}{2n^2 + 1} - \frac{1}{2} &< \varepsilon \\ -\varepsilon < \frac{2n^2 + 4 - 2n^2 - 1}{4n^2 + 2} &< \varepsilon \\ -\varepsilon < \frac{3}{4n^2 + 2} &< \varepsilon \end{aligned}$$

In quest'ultima disuguaglianza, come prima, il primo termine è sempre vero, mentre il secondo è da risolvere.

$$\begin{aligned} \frac{3 - 4\varepsilon n^2 - 2\varepsilon}{4n^2 + 2} &< 0 \\ 4\varepsilon n^2 + 2\varepsilon - 3 &> 0 \\ 4\varepsilon n^2 &> 3 - 2\varepsilon \\ n^2 &> \frac{3 - 2\varepsilon}{4\varepsilon} \\ n &> \sqrt{\frac{3 - 2\varepsilon}{4\varepsilon}} \end{aligned}$$

Ed anche in questo caso abbiamo trovato la soluzione che cercavamo per n_ε.

Tuttavia, si devono ancora considerare un paio di casi sul comportamento della successione a_n quando si considerano termini sempre più avanti. Questi infatti possono, ad esempio, diventare sempre più grandi (in positivo, o negativo). In questo caso scriveremo:

$$\lim_{n \to \infty} a_n = +\infty$$

oppure

$$\lim_{n \to \infty} a_n = -\infty$$

In questo caso le definizioni di limite cambieranno leggermente: infatti ora non dobbiamo più chiedere che gli a_n siano compresi in un certo intervallo, ma che siano sempre superiori al ε arbitrario maggiore di zero, o sempre minore di ε, questa volta considerato negativo (oppure positivo, ma con segno cambiato nella disuguaglianza):

- $\lim_{n\to\infty} a_n = +\infty \stackrel{def}{=} \forall \varepsilon > 0 : \exists n_\varepsilon \in \mathbb{N} : a_{n>n_\varepsilon} > \varepsilon$
- $\lim_{n\to\infty} a_n = +\infty \stackrel{def}{=} \forall \varepsilon < 0 : \exists n_\varepsilon \in \mathbb{N} : a_{n>n_\varepsilon} < \varepsilon$

I tre limiti $l \in \mathbb{R}, +\infty, -\infty$ sono esclusivi l'uno dell'altro, ma ciò non significa che una successione debba obbligatoriamente tendere ad uno di questi tre valori.

Se una successione tende ad un valore reale, si dice allora che la successione è *convergente*, se tende a più o meno infinito che è *divergente*, altrimenti che è *irregolare* o *indeterminata*, il che significa che non tende ad alcun valore.

Esempio 4.1.3 *Stabilire che la successione*

$$a_n = (-1)^n \frac{n}{n+1}, n \geq 1$$

ha carattere irregolare.

L'unico sistema che abbiamo per verificare ciò è mostrare che non può tendere nè ad un valore finito, nè a $+\infty$, nè a $-\infty$. Inanzitutto osserviamo che

$$|a_n| = \frac{n}{n+1}$$

E questi sono tutti valori minori di 1. Quindi:

$$-1 < a_n < 1$$

I suoi termini sono alternativamente positivi e negativi (dovuto ciò al termine moltiplicativo $(-1)^n$). Quindi, sicuramente non tende nè a $+\infty$, nè a $-\infty$. Mostriamo che, se a_n tende ad un valore reale, questo valore deve essere 0, per mezzo di una breve dimostrazione per assurdo.

Infatti, se tendesse ad un $l > 0$, si potrebbe prendere sicuramente un ε tale che $0 < l - \varepsilon$. In questo caso, l'intervallo $]l - \varepsilon, l + \varepsilon[$ sarebbe composto solo da valori positivi, e la succesione a_n dovrebbe, da un certo n_ε in poi, essere costituita solo da termini positivi, cosa che non è possibile per la natura alterna dei suoi segni. Quindi non può essere vero che $l > 0$. Analogo ragionamento per $l < 0$. (Vedi più avanti il Teorema sulla permanenza del segno – 4.1.2).

Quindi ora rimane da verificare se a_n tende a 0 o meno. Utilizzando la definizione di limite:

$$\stackrel{?}{\exists}\, n_\varepsilon : \forall n > n_\varepsilon : \left|(-1)^n \frac{n}{n+1}\right| < \varepsilon$$

Possiamo ora compiere queste semplici operazioni:

$$\frac{n}{n+1} < \varepsilon$$

$$n < n\varepsilon + \varepsilon$$

$$\underbrace{n(1-\varepsilon)}_{a} < \underbrace{\varepsilon}_{b}$$

Ora, se $\varepsilon > 1$, allora $a < b$, e la disuguaglianza è sempre verificata, per qualunque n.
Se invece $0 < \varepsilon < 1$, allora devo risolvere:

$$n < \frac{\varepsilon}{1-\varepsilon}$$

Quindi abbiamo che non per tutti gli ε si trova un n_ε - allora, a_n non tende neanche a 0.

Avendo esaurito tutte le possibilità, l'ultima che rimane è che la successione data sia quindi irregolare.

Esempio 4.1.4 *Data*

$$a_n = \alpha^n, \alpha > 1$$

Dimostrare che $a_n \to +\infty$.

La dimostrazione è molto semplice. Utilizzando la definizione:

$$\forall \varepsilon > 0 : \alpha^n > \varepsilon$$

Applichiamo il logaritmo in base α da entrambi i membri (questa volta non si deve invertire il verso perché, per basi maggiori di 1, il logaritmo è una funziona crescente*).*

$$n > \ln_\alpha \varepsilon$$

Quindi abbiamo trovato il nostro n_ε.

La piccola dimostrazione per assurdo utilizzata nell'esempio 4.1.3 si può facilmente essere estesa al seguente teorema:

Teorema 4.1.2 (Teorema della permanenza del segno) *Data una successione a_n che tende ad un valore finito l, questa successione ha* definitivamente *il segno di l.*

Con *definitivamente* si intenderà sempre da ora in poi il concetto "da un certo n in poi".

4.1.3 Successioni monotòne

Data una successione a_n questa si dirà:

- Monotòna crescente se $a_n \leq a_{n+1}$
- Monotòna decrescente se $a_n \geq a_{n+1}$
- Strettamente crescente se $a_n < a_{n+1}$
- Strettamente decrescente se $a_n > a_{n+1}$

In pratica, una successione è monotona se mantiene sempre lo stesso "andamento" (crescente o descrescente) rimanendo al massimo per alcuni o tutti i valori costante, è strettamente crescente o decrescente se quest'ultima opzione non è consentita.

Queste successioni sono di particolare importanza, infatti per essere vale il seguente:

Teorema 4.1.3 (Teorema sull'esistenza del limite per successioni monotone) *Una successione a_n definitivamente crescente (decrescente) ammette sempre limite, e questo coincide con l'estremo superiore (inferiore) dell'insieme $\{a_n \mid n \in \mathbb{N}\}$.*

In questo caso si usa una definizione più estesa del solito per "estremo superiore" (e analogamente inferiore): ovvero si considera anche il caso in cui l'insieme non sia limitato: in tal caso si prende ∞ come estremo. La simbologia per indicare ciò è:

$$a_n \uparrow l \text{ oppure } a_n \downarrow l$$

Questo teorema vale solo se l'ambiente considerato è $\mathbb{R}$.

Dimostriamo questo enunciato per l finito. In pratica quello che dobbiamo mostrare è che, se

$$l = supA, A = \{a_n \mid n \in \mathbb{N}\}$$

e

$$a_n \leq a_{n+1}$$

allora

$$|a_n - l| < \varepsilon,\ \forall n > n_\varepsilon$$

(definizione di limite), il che equivale a dire che

$$\begin{cases} l - \varepsilon < a_n \\ a_n < l - \varepsilon \end{cases},\ \forall n > n_\varepsilon$$

La seconda disuguaglianza è facilmente dimostrabile: dato che l per ipotesi è l'estremo superiore finito di A, allora $\forall a_n \in A, a_n \leq l$, quindi $\forall a_n \in A, a_n < l + \varepsilon$. Allora questo membro non solo è vero per ogni $n > n_\varepsilon$, ma addirittura per ogni n.

Per quanto riguarda la prima disuguaglianza inanzitutto osserviamo che $l = \min \mathsf{M}_A$, e quindi $l - \varepsilon \notin \mathsf{M}_A$, il che è la stessa cosa che dire: non è vero che $l - \varepsilon > a_n, n \in \mathbb{N}$. Quindi, $\exists \bar{n} \mid a_{\bar{n}} > l - \varepsilon$. Allora, dato che $a_{n+1} \geq a_n$, questa proprietà sarà vera anche per tutti gli $n > \bar{n}$ – a questo punto basta prendere $n_\varepsilon = \bar{n}$, per ottenere la prima parte della disugualianza. Ecco quindi dimostrata la proposizione per $l \in \mathbb{R}$.

Invece per $l = +\infty$, abbiamo come ipotesi il fatto che la successione sia crescente e che l'insieme non sia superiormente limitato, ovvero che per ogni $\varepsilon > 0$, esiste sempre almeno un a_n maggiore di esso nell'insieme. Espresso simbolicamente ciò si traduce nelle seguenti condizioni:

$$\nexists \sup\{a_n \mid n \in \mathbb{B}\} \Rightarrow \forall \varepsilon > 0 : \exists a_n : a_n > \varepsilon$$

$$a_n \leq a_{n+1}$$

E quello che vogliamo dimostrare è che

$$\forall \varepsilon > 0, \exists n_\varepsilon : a_n > \varepsilon, \forall n > n_\varepsilon$$

Ora, dalla prima proposizione, siamo già sicuri che esiste almeno un a_n con la proprietà sopra descritta (chiamiamolo $a_{\bar{n}}$). Ma, dato che $a_{n+1} \geq a_n$, allora questa proprietà varrà sicuramente anche per tutti gli a_n successivi. Ecco dimostrato il teorema anche per $l = +\infty$. Ovviamente per il caso di una successione descrescente il discorso è il medesimo.

4.1.4 Regole algebriche dei limiti

Si possono dimostrare facilmente una serie di proprietà alquanto intuitive sull'utilizzo di regole algebriche ai limiti.

Si prendano due successioni a_n e b_n, tendenti rispettivamente ad a e b. Abbiamo allora le seguenti:

$$\begin{aligned} a_n \pm b_n &\to a \pm b \\ a_n b_n &\to ab \\ \frac{a_n}{b_n} &\to \frac{a}{b} \quad (b_n, b \neq 0) \\ a_n^{b_n} &\to a^b \quad (a_n, a > 0) \end{aligned}$$

Un caso delicato è quell nella forma $a/0$. Infatti, di solito non si può dire nulla di questa forma (vedi sotto quello che si dice riguardo alle *forme indeterminate*), tuttavia se si sa che la successione tende a 0, lo fa con tutti valore positivi diversi da 0, allora il risultato è infinito, col segno di a. Similmente se la successione tende a 0 con tutti valori negativi diversi da 0 (ma il segno dell'infinito sarà cambiato).

Un utile teorema per il calcolo è il seguente:

Teorema 4.1.4 (Teorema del confronto per le successioni) $a_n \leq b_n \leq c_n \ \wedge \ a_n \to l, c_n \to l \ \Rightarrow \ b_n \to l$

Ovvero: se una successione è "costretta" tra altre due, e quelle due tendono ad un certo valore, anche la prima lo farà.

Se invece a o b sono valori infiniti invece che finiti? Allora possiamo anche in questo caso utilizzare l'intuito e trovare effettivamente le giuste regole, che ci consentono di scrivere risultati di operazioni che coinvolgono il segno ∞. Si badi bene che le seguenti scritture non sono *formalmente* corrette, tuttavia esprimono in modo semplice le regole che governano questo genere di limiti.

$$
\begin{aligned}
a + \infty &= \infty \\
a - \infty &= -\infty \\
+\infty + \infty &= +\infty \\
-\infty - \infty &= -\infty \\
a \cdot (+\infty) &= \begin{cases} +\infty \text{ se } a > 0 \\ -\infty \text{ se } a < 0 \end{cases} \\
a \cdot (-\infty) &= \begin{cases} -\infty \text{ se } a > 0 \\ +\infty \text{ se } a < 0 \end{cases} \\
(+\infty) \cdot (+\infty) &= (-\infty) \cdot (-\infty) = +\infty \\
(+\infty) \cdot (-\infty) &= -\infty \\
\frac{a}{\pm\infty} &= 0
\end{aligned}
$$

Si vede tuttavia che alcuni casi non sono stati considerati. Questi casi mancanti sono detti *forme indeterminate* e su di essere non si può dire nulla a priori, senza compiere uno studio specifico per la successione considerata. Queste sono:

- $\infty - \infty$
- $0 \cdot (\pm\infty)$
- $0/0$
- ∞/∞

Facciamo alcuni esempi per convincerci della molteplicità di risultati che possono avere tali forme.

Esempio 4.1.5 *Date*

$$a_n = n^2, \quad b_n = \frac{1}{n}$$

Trovare $\lim_{n\to+\infty} a_n b_n$.

È facile verificare che $a_n \to +\infty$ *e* $b_n \to 0$. *Ci troviamo quindi nel caso* $0 \cdot +\infty$. *Il risultato che si ricava in questo caso è:*

$$a_n b_n = n^2 \cdot \frac{1}{n} = \frac{n^2}{n} = n$$

Che ovviamente tende a $+\infty$.

Esempio 4.1.6 *Date*

$$a_n = \frac{1}{n^2}, \quad b_n = n$$

Trovare $\lim_{n\to+\infty} a_n b_n$.

Anche questo è un caso di $0 \cdot +\infty$. *Qui invece il risultato è:*

$$a_n b_n = n \cdot \frac{1}{n^2} = \frac{n}{n^2} = \frac{1}{n}$$

Che tende invece questa volta a 0.

Per quanto riguarda il quoziente la soluzione è ancora più delicata. La forma "sicura" su cui riusciamo a trattare è

$$\frac{1}{\pm\infty} \to 0$$

Mentre

$$\frac{1}{0} \not\to \infty$$

Infatti, già il fatto che non si possa stabilire a priori il segno dell'infinito dovrebbe essere d'allarme. Infatti, questo è dovuto al modo con cui la successione al denominatore di sinistra si avvicina a zero: è vero infatti che se lo fa con tutti termini positivi, il suo reciproco tenderà a $+\infty$, viceversa per termini negativi, ma potrebbe anche tendere a 0 con segni alterni, o comunque non definitivamente uguali, ed in questo caso la successione sarebbe *irregolare*.

Esempio 4.1.7 *Trovare il carattere della successione*

$$a_n = \frac{n}{(-1)^n}$$

Possiamo osservare inanzitutto che questa successione è il reciproco di:

$$a_n = \frac{(-1)^n}{n}$$

Che può essere calcolata ricadendo sotto il caso $1/\infty$: infatti tende a 0. Tuttavia il suo reciproco NON *tende a $\pm\infty$! Troviamo infatti che i termini sono alternativamente di segno positivo e negativo, e sempre più grandi in valore assoluto. Quindi, ovviamente non hanno alcun valore a cui si avvicinano – la successione considerata è* irregolare.

Le regole algebriche appena trovate semplificano molto spesso il calcolo di limiti.

Esempio 4.1.8 *Trovare il limite della successione*

$$a_n = \frac{n^2+2}{2n^2+1}$$

Applichiamo una qualche manipolazione sulla regola della successione:

$$a_n = \frac{n^2\left(1+2/n^2\right)}{n^2\left(2+1/n^2\right)} = \frac{1+2/n^2}{2+1/n^2}$$

Ora, abbiamo che $2/n^2$ e $1/n^2$ tendono a 0, essendo nella forma $1/\infty$. Sommati quindi a delle "successioni" (1 e 2, rispettivamente) a termini costanti (1 e 2, per qualunque n ovviamente), daranno come risultato 1 e 2, e il loro rapporto è $1/2$.

Esempio 4.1.9 *Trovare il limite della successione*

$$a_n = \frac{n^2+n-1}{n^4-3n^3+12}$$

Con passaggi simili ai precedenti, si arriva questa volta alla forma:

$$\frac{1+1/n-1/n^2}{n^2\left(1-3/n+12/n^4\right)}$$

Che, come è facile vedere, è riconducibile questa volta alla forma $1/\infty$, e che quindi da 0 come risultato.

Si può quindi estrapolare da questo paio di esempi una regola più generale per il rapporto di due polinomi (per la quale tuttavia conviene di più imparare l'idea che vi sta alla base piuttosto che la regola mnemonica) e che è:

$$\lim_{n\to\infty} \frac{p_0+p_1n+p_2n^2+\cdots+p_kn^k}{q_0+q_1n+q_2n^2+\cdots+q_hn^h} = \begin{cases} \infty \text{ col segno di } p_k/q_h \text{ se } k>h \\ 0 \text{ se } k<h \\ p_k/q_h \text{ se } k=h \end{cases}$$

Esempio 4.1.10 *Calcolare il valore a cui converge la successione*

$$\left(\sqrt{n+2}-\sqrt{n+1}\right)$$

Anche in questo caso dobbiamo compiere un "raccoglimento":

$$\frac{n+2-n-1}{\sqrt{n+2}+\sqrt{n+1}} = \frac{1}{\sqrt{n+2}+\sqrt{n+1}} \to 0$$

4.1.5 Il numero di Nepero

Lo studio della successione

$$a_n = \left(a + \frac{1}{n}\right)^n$$

È particolarmente importante, perché conduce alla definizione del *numero di Nepero*, indicato con la lettera e, un valore di importanza fondamentale nell'Analisi insieme a π.

In questa sede, con gli strumenti che abbiamo, possiamo dimostrare che la successione data ha limite, che questo è reale, ed è compreso tra 2 e 4.

Inanzitutto, mostriamo che la successione presa in questione è strettamente crescente. Per fare ciò, dobbiamo dimostrare che:

$$a_n < a_{n+1}$$

che è come dire

$$a_{n-1} < a_n$$

ovvero che

$$\left(1 + \frac{1}{n-1}\right)^{n-1} < \left(1 + \frac{1}{n}\right)^n$$

Nel nostro caso prenderemo $n \geq 2$. In effetti, ai nostri fini, ci interessa solo che la successione sia *definitivamente* crescente.

Un modo per dimostrare ciò è quello di calcolare il rapporto tra il secondo e il primo membro e mostrare che questo è maggiore di 1:

$$\frac{(1 + 1/n)^n}{[1 + 1/(n-1)]^{n-1}} \overset{?}{>} 1$$

Questo perché una frazione è maggiore di 1 solo se il suo numeratore è superiore al denominatore. Ora sviluppiamo questa frazione:

$$\begin{aligned}
&\frac{[(n+1)/n]^n}{[1 + 1/(n-1)]^n \, [1 + 1/(n-1)]^{-1}} \\
=&\frac{[(n+1)/n]^n}{[n/(n-1)]^n \, [n/(n-1)]^{-1}} \\
=&\frac{\left[(n^2-1)/n^2\right]^n}{1 - 1/n} \\
=&\frac{\left(1 - 1/n^2\right)^n}{1 - 1/n}
\end{aligned}$$

Adesso dobbiamo applicare la cosiddetta *disuguaglianza da J. Bernoulli*, che dice questo:

Teorema 4.1.5 (Disuguaglianza di Bernoulli) $t > -1, t \neq 0, n \geq 2 \Rightarrow (1+t)^n > 1 + nt$

Applicata al numeratore della frazione così trovata, poniamo $t = -1/n^2$, ed n proprio uguale ad n. $n \geq 2$, per ipotesi. t così definita invece ha valori che partono da $-1/4$, per $n = 2$, e crescono, quindi senz'altro soddisferà le condizioni della disuguaglianza di Bernoulli. Se la applichiamo troviamo quindi:

$$\left(1 - \frac{1}{n^2}\right)^n > 1 - n\frac{1}{n^2} = 1 - \frac{1}{n}$$

Che è proprio il denominatore della frazione. Quindi, dato che il numeratore è maggiore del denominatore, la frazione stessa è maggiore di 1, e così abbiamo dimostrato la disuguaglianza di partenza. La successione data è quindi strettamente crescente.

Ci rimane ora da dimostrare che è limitata. Osserviamo inanzitutto che, se la successione a_n è strettamente crescente, questo significa che tutti i suoi termini sono maggiori o uguali del primo. a_1 è facile da calcolare, e risulta 2, quindi abbiamo che

$$2 \leq a_n$$

Il limite inferiore della successione l'abbiamo quindi trovato. Ora prendiamo una nuova successione che ci verrà in aiuto:

$$b_n = \left(1 + \frac{1}{n}\right)^{n+1}$$

Notiamo che l'unica differenza da a_n è l'esponente che è maggiore di 1. Confrontandola con la a_n, troviamo che senz'altro

$$a_n \leq b_n$$

Proprio a causa dell'esponente maggiore. Tuttavia, applicando passaggi del tutto analoghi ai precedenti, si trova che b_n è strettamente *decrescente*. Quindi, tutti i suoi termini saranno *minori* di 4, che è il suo primo termine.

$$b_n \leq 4$$

Unendo le varie disuguaglianze ottenute, troviamo:

$$2 \leq a_n \leq b_n \leq 4$$

Ovvero

$$2 \leq a_n \leq 4$$

Abbiamo quindi che la successione a_n è limitata. Per il teorema sulle successioni monotone, abbiamo quindi che a_n, essendo crescente e limitata, ha limite finito. Questo limite, che sarà sicuramente compreso tra 2 e 4, viene indicato, come detto ad inizio paragrafo, con e e viene chiamato *numero di Nepero*.

In realtà si possono anche dimostrare altre due proprietà, ovvero che non solo $2 \leq e \leq 4$, ma che $2 < e < 3$ e che e è un numero trascendente (ovvero che non esiste alcuna equazione polinomiale a coefficienti interi che abbia e tra le sue radici).

4.1.6 Infinitesimi ed infiniti

Una successione viene chiamata *infinitesimo* se tende a 0:

$$a_n \to 0$$

Analogamente, una successione viene chiamata *infinito* se tenda a ∞:

$$a_n \to \infty$$

Tuttavia, presi, ad esempio, due infinitesimi, si potrebbe voler sapere quale tende a 0 più velocemente (ovvero quale infinitesimo è *di ordine superiore* rispetto all'altro, cosa che vedremo essere alquanto utile). Questo confronto si può compiere analizzando il limite del loro rapporto (il quale ovviamente risulta sempre a prima vista una forma indeterminata):

$$\frac{a_n}{b_n} \to \begin{cases} 0 \to a_n \text{ è } \textit{di ordine superiore} \text{ rispetto a } b_n \\ \pm\infty \to a_n \text{ è } \textit{di ordine inferiore} \text{ rispetto a } b_n \\ l \neq 0 \to a_n \text{ è dello } \textit{stesso ordine} \text{ di } b_n \end{cases}$$

Il ragionamento che sta alla base è abbastanza semplice: per fare un esempio, prendiamo $a_n = 1/n^2$ e $b_n = 1/n$. Entrambi tendono a 0 evidentemente, ma a_n lo fa più velocemente, infatti osservando i loro valori per alcuni n crescenti abbiamo:

n	a_n	b_n
1	1	1
2	1/4	1/2
3	1/9	1/3
...	...	...
100	1/10000	1/100
...	...	...

Quindi è ovvio aspettarsi che il rapporto a_n/b_n sia un numero sempre più piccolo: sebbene il denominatore diminuisca, e quindi faccia tendere la frazione a $+\infty$, il numeratore va a 0 *più velocemente*, prendendo quindi il sopravvento e facendo tendere la frazione complessivamente a 0.

n	a_n/b_n
1	1
2	1/2
3	1/3
...	...
100	1/100
...	...

Prendiamo il caso invece di infinitesimi dello stesso ordine: ad esempio $a_n = 1/n$ e $b_n = 3/n$. Ad occhio si vede subito che il secondo infinitesimo è per ogni n valido minore del secondo. Tuttavia, se si applica la regola usata, si ha che il limite del loro rapporto è $l = 1/3 \neq 0$, e quindi sono dello stesso ordine. Questo mostra che il confronto tra infinitesimi non è un semplice "minore", "maggiore" o "uguale", bensì coglie distinzioni più forti: due serie che differiscono per una costante moltiplicativa sono considerati equivalenti, e solo se divergono in modo più marcato (come l'esempio precedente) vengono considerati di ordini diversi.

Confronti analoghi a quelli tra gli infinitesimi si possono applicare tra infiniti:

$$\frac{a_n}{b_n} \to \begin{cases} 0 \to a_n \text{ è } \textit{di ordine inferiore} \text{ rispetto a } b_n \\ \pm\infty \to a_n \text{ è } \textit{di ordine superiore} \text{ rispetto a } b_n \\ l \neq 0 \to a_n \text{ è dello } \textit{stesso ordine} \text{ di } b_n \end{cases}$$

In ogni caso, se il limite di tale rapporto non esiste, si dice che gli infiniti o infinitesimi presi in questione *non sono confrontabili.*

Ci sono alcuni infiniti il cui confronto è basilare, ma di cui non daremo la dimostrazione perché per ora non ne possediamo gli strumenti, e sono:

$$\frac{\alpha^n}{n^k} \to \infty, \quad \alpha > 1, k > 0$$

Ovvero qualunque esponenziale è sempre un infinito di ordine superiore rispetto ad una potenza.

$$\frac{n^k}{\ln_\alpha n} \to \infty, \quad \alpha > 1, k > 0$$

Ovvero qualunque potenza è sempre un infinito di ordine superiore rispetto ad un logaritmo (a dire il vero si potrebbe mostrare anche che qualunque potenza è un infinito di ordine superiore rispetto ad un polilogaritmo, ovvero un polinomio che ha per termini potenze di logaritmi!).

Possiamo quindi instaurare una specie di "gerarchia" di infiniti, dal più lento al più veloce (ovvero da quello d'ordine inferiore a quello d'ordine superiore):

$$\ln_\alpha n,\ n^k,\ \alpha^k$$

A questi possono essere aggiunti anche altre tre importanti successione che vedremo nell'esercizio 4.1.11.

Nel caso di due infinitesimi o infiniti dello stesso ordine, si distingue il caso particolarmente importante nel quale $a_n/b_n \to 1$. In questo caso si usa dire che le due successioni a_n e b_n sono *asintotiche* e si indica con la scrittura:

$$a_n \sim b_n$$

che si legge "a_n è *asintotico* a b_n". Spesso si compie una trasformazione tra due successioni che sono dello stesso ordine per renderle asintotiche, infatti:

$$\lim_{n\to\infty} \frac{a_n}{b_n} = l \neq 0,\ l \in \mathbb{R}$$

porta a:

$$\lim_{n\to\infty} \frac{a_n/l}{b_n} = l/l = 1 \implies \frac{a_n}{l} \sim b_n$$

ed anche

$$\lim_{n\to\infty} \frac{a_n}{b_n \cdot l} = l/l = 1 \implies a_n \sim l \cdot b_n$$

Questo si fa perché spesso, nelle applicazioni delle proprietà delle successioni asintotiche (come quella che vedremo di seguito e varie altre) è facile portare dentro o fuori le formule un fattore costante.

Si dimostra facilmente che:

- Se $a_n \sim b_n$ le due successioni hanno lo stesso comportamento per $n \to +\infty$, ovvero convergono allo stesso limite, o entrambe divergono a più o meno infinito, o entrambe sono irregolari.
- L'operatore "$\sim$" gode della proprietà transitiva ($a_n \sim b_n, b_n \sim c_n \Rightarrow a_n \sim c_n$).
- Si può sostituire in un prodotto o rapporto più successioni con delle successioni ad esse asintotiche - ad es:

  $$a_n \sim a'_n, b_n \sim b'_n, c_n \sim c'_n \Rightarrow \frac{a_n b_n}{c_n} = \frac{a'_n b'_n}{c'_n}$$

  (da notare che lo stesso discorso *non* vale per somme, differenze o esponenziali)
- Un comodo modo per dimostrare che $a_n \sim b_n$ è dimostrare che $a_n = b_n c_n, c_n \to 1$ – che le due proposizioni siano equivalenti si ricava semplicemente dividendo entrambi i membri della seconda per b_n.

4.1.7 Alcune importanti proprietà dei limiti

In questo paragrafo presenteremo alcuni importanti teoremi.

Criterio 4.1.6 (Criterio del rapporto (per le successioni)) *Data una successione a_n, in base al limite della successione a_{n+1}/a_n si può dire che:*

$$\lim_{n\to+\infty} \frac{a_{n+1}}{a_n} = \begin{cases} 0 \leq l < 1 \ \Rightarrow\ a_n \text{ converge a } 0 \\ l > 1, l = +\infty \ \Rightarrow\ a_n \text{ diverge a } +\infty \\ l = 1 \ \Rightarrow\ \text{nulla si può dire sul risultato} \end{cases}$$

Questo criterio da un grande aiuto spesso per la determinazioni dei limiti a 0 o $+\infty$. Tuttavia possiede il punto debole che, nel frequente caso in cui il limite del rapporto risulti 1, esso non fornisce alcuna informazione: il risultato potrebbe essere qualunque. Tra gli esempi riguardo alla gerarchia dei limiti troveremo invece alcuni esempi di applicazione con successo di tale criterio.

Criterio 4.1.7 (Criterio della radice (per le successioni)) *Data una successione a_n positiva ($a_n > 0, \forall n$), si ha che, se:*

$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = l \in \mathbb{R}$$

allora

$$\lim_{n\to\infty} \sqrt[n]{a_n} = \sqrt[n]{l}$$

Questo criterio, dalla forma leggermente diversa da quello di prima, è utile solo per calcolare le successioni composte da radici n-sime, come ad esempio per dimostrare che $\sqrt[n]{\alpha} \to 1, \alpha > 0$, oppure che $\sqrt[n]{n} \to 1$.

Teorema 4.1.8 (Teorema del confronto) *Se si hanno tre successioni:*

$$a_n \leq b_n \leq c_n, \forall n \in \mathbb{N}$$

e se $a_n \to l$ e $c_n \to l$, allora $b_n \to l$.

Questo teorema in pratica esprime il fatto che, se due successione "circondano" una terza successione, e se queste due successione per n che tende ad infinito, tendono ad uno stesso valore, l'ultima successione è "costretta" ad assumere anche lei lo stesso valore.

La dimostrazione è alquanto semplice. Per definizione di limite abbiamo che:

$$l - \varepsilon < a_n < l + \varepsilon$$

e

$$l - \varepsilon < c_n < l + \varepsilon$$

Ma d'altronde, per ipotesi

$$a_n \leq b_n \leq c_n$$

unendo le disuguaglianze risulta quindi

$$l - \varepsilon < a_n \leq b_n \leq c_n < l + \varepsilon$$

ovvero

$$l - \varepsilon < b_n < l + \varepsilon$$

che è la definizione di limite in l per b_n. CVD.

Infine, ricordiamo di nuovo il Teorema di permanenza del segno (4.1.2) ed anche i vari teoremi sull'aritmetizzazione parziale dell'infinito e sulle operazioni tra limiti.

Esempio 4.1.11 *Dimostrare che la seguente sequenza di infiniti è da quello di ordine inferiore a quello di ordine superiore:*

$$\ln_\alpha n,\ n^k,\ \alpha^k, n!, n^n$$

Per quanto riguarda i primi tre infiniti, ne è già stata assicurata la veridicità nella parte teorica. Rimangono quindi ancora due limiti da considerare.

$$\frac{\alpha^n}{n!} \to ?$$

Utilizziamo il criterio del rapporto:

$$\frac{a_{n+1}}{a_n} = \frac{\alpha^{n+1}}{(n+1)!} \cdot \frac{n!}{\alpha^n} = \frac{\alpha \cdot \alpha^n}{(n+1)n!} \cdot \frac{n!}{\alpha^n} = \frac{\alpha}{n+1} \to 0$$

Il limite cercato quindi è 0: $n!$ è effettivamente d'ordine superiore rispetto a α^n.

$$\frac{n!}{n^n} \to ?$$

Utilizzando sempre il criterio del rapporto:

$$\frac{a_{n+1}}{a_n} =$$

$$\frac{(n+1)!}{(n+1)^{n+1}} \cdot \frac{n^n}{n!} =$$

$$\frac{(n+1)n!}{(n+1)(n+1)^n} \cdot \frac{n^n}{n!} =$$

$$\left(\frac{n}{n+1}\right)^n =$$

$$\frac{1}{\left(1+\frac{1}{n}\right)^n} \to \frac{1}{e}$$

Quindi per il Teorema del rapporto, anche in questo caso il limite cercato tende a 0: n^n è effettivamente d'ordine superiore rispetto a $n!$.

Esempio 4.1.12 *Calcolare il limite di*

$$a_n = \sqrt[n]{n^{300}}$$

Possiamo trasformare a_n *in*

$$a_n = n^{\frac{300}{n}}$$

Prendendo il logaritmo naturale della successione troviamo:

$$\ln a_n = \frac{300}{n} \ln n = 300 \frac{\ln n}{n}$$

Ma noi sappiamo dalla gerarchia degli infiniti che $\ln n$ *è un infinito d'ordine inferiore a* n*, e quindi* $\ln n/n \to 0$*. Quindi:*

$$\ln a_n \to 300 \cdot 0 = 0$$

Il che vuol dire che

$$a_n = e^{\ln a_n} \to e^0 = 1$$

Esempio 4.1.13 *Calcolare il limite di*

$$a_n = \left(1 + \frac{1}{n^2}\right)^n$$

Come prima:

$$\ln a_n = \ln \left(1 + \frac{1}{n^2}\right)^n = n \ln \left(1 + \frac{1}{n^2}\right) = \frac{n^2 \ln \left(1 + \frac{1}{n^2}\right)}{n} = \frac{\ln \left(1 + \frac{1}{n^2}\right)^{n^2}}{n}$$

Ma il valore all'interno del logaritmo tende ad e*, quindi l'intera espressione tende a:*

$$\frac{\ln e}{n} \to 0$$

E tornando alla successione originaria:

$$a_n = e^{\ln a_n} \to e^0 = 1$$

4.2 Serie numeriche

4.2.1 Definizione di serie

Ad un livello puramente intuitivo, una serie è la somma di infiniti termini. In pratica, può essere vista come la somma di tutti i termini di una successione, quindi una volta individuata la legge della successione si è determinata anche la serie corrispondente.

Tuttavia, non abbiamo mezzi diretti per esprimere la somma di *infiniti* termini. Quindi, si deve utilizzare un altro metodo. Ad esempio, se vogliamo definire la serie composta dai termini della successione a_n, potremmo utilizzare questo modo: prima definiamo una nuova successione, chiamata *successione delle somme parziali*, cosìdefinita:

$$\begin{aligned} s_0 &= a_0 \\ s_1 &= a_0 + a_1 \\ s_2 &= a_0 + a_1 + a_2 \\ &\dots \\ s_n &= a_0 + a_1 + \dots + a_n \end{aligned}$$

Com'è facile intuire, la serie degli a_n sarà il limite delle somme parziali, per n che tende ad infinito. Quindi, la serie di termini a_n viene definita convergente ad un valore reale, divergente a più o meno infinito o è irregolare a seconda del comportamento della successione delle somme parziali della serie stessa.

La simbologia utilizzata per indicare una serie è di solito

$$a_0 + a_1 + \dots + a_n + \dots$$

o meglio

$$\sum_{n=0}^{\infty} a_n$$

Quindi, la definizione data prima si può indicare come:

$$\sum_{n=0}^{\infty} a_n = A \iff \lim_{n \to \infty} s_n = A$$

Talvolta non si parte da $n = 0$ ma da un $n = N$ - ovviamente il discorso fatto rimane analogo.

Esempio 4.2.1 *Determinare il carattere della serie geometrica, ovvero di:*

$$1 + q + q^2 + q^3 + \cdots + q^n + \cdots = \sum_{n=0}^{\infty} q^n, \quad q \in \mathbb{R}$$

Dobbiamo per prima cosa trovare la legge delle somme parziali. Procediamo così: calcoliamo inanzitutto s_n:

$$s_n = 1 + q + q^2 + \cdots + q^n$$

Quindi moltiplichiamo per q:

$$q \cdot s_n = q + q^2 + q^3 + \cdots + q^{n+1}$$

ed ora calcoliamo la differenza tra i due membri:

$$s_n - qs_n = (1 + q + q^2 + q^3 + \cdots + q^n) - (q + q^2 + q^3 + \cdots + q^{n+1})$$

$$(1 - q)s_n = 1 - q^{n+1}$$

$$s_n = \frac{1 - q^{n+1}}{1 - q}$$

Adesso che abbiamo un'espressione generale per l'n-simo termine della successione delle somme parziali (ottenuta presupponendo, si osservi, che $q \neq 1$) possiamo calcolare il suo limite per n che tende ad infinito. Considerando i vari casi, otteniamo:

- *Per $|q| < 1$, q^{n+1} tende a 0, quindi la serie tende al valore $1/(1 - q)$.*
- *Per q=1, esaminiamo direttamente la serie (non possiamo utilizzare la formula come premesso perché è stata ottenuta partendo dal presupposto che $q \neq 1$): diventa:*

  $$1 + 1 + 1 + \cdots + 1 + \ldots$$

  che tende chiaramente ad infinito
- *Per $q > 1$, q^{n+1} tende ad infinito, quindi abbiamo che il limite tende anch'esso ad infinito.*
- *Per $q \leq -1$ la potenza q^{n+1} diventa una successione irregolare, e cosìrisulta anche il limite della successione delle somme parziali.*

Riassumendo:

$$\sum_{n=0}^{\infty} q^n = \begin{cases} \dfrac{1}{1-q} & \text{se } |q| < 1 \\ +\infty & \text{se } q \geq 1 \\ \text{non esiste} & \text{se } q \leq -1 \end{cases}$$

4.2.2 Proprietà delle serie

Alcune serie, come quella geometrica qui discussa, sono particolarmente importanti. Un altro genere di serie già analizzato è quello della *serie armonica*, ovvero:

$$1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} + \cdots = \sum_{n=1}^{\infty} \frac{1}{n}$$

che abbiamo dimostrato divergere a più infinito. D'altronde, però, se consideriamo più genericamente

$$\sum_{n=1}^{\infty} \frac{1}{n^k}$$

(ovvero la cosiddetta "serie armonica generalizzata") cosa otteniamo? Ora i valori della serie decrescono molto più rapidamente di prima, quindi potrebbe succedere che la loro somma sia un numero reale. Si può dimostrare che *la serie armonica generalizzata converge per $k > 1$ e diverge per i $k \leq 1$*. Ad esempio, per $k = 2$, converge al valore $\pi^2/6$. Intanto possiamo già dimostrare che essa è convergente utilizzando un paio di considerazioni.

Prima però enunciamo una facile proposizione:

Teorema 4.2.1

$$\sum_{n=0}^{\infty} a_n \Longrightarrow a_n \to 0$$

Importante è il fatto che l'implicazione va in un unico verso: ovvero, la convergenza a 0 della successione dei termini della serie è condizione *necessaria*, ma non *sufficiente*! Questo significa che si può dimostrare che una serie non è convergente mostrando che la successione a_n non converge a 0, ma il fatto che a_n tenda a 0 non permette di dire che la serie sia convergente (ad esempio la serie armonica semplice ha termini che tendono a 0, ma non converge).

Questa proposizione si dimostra facilmente: infatti, se supponiamo che la serie in questione sia convergente ad s, significa che $s_n \to s$, e quindi:

$$a_n = (s_n - s_{n+1}) \to (s - s) = 0$$

CVD.

D'ora in poi, a meno che non venga detto il contrario, si considereranno solo serie a termini positivi, ovvero in cui $a_n \geq 0$. Qual'è il vantaggio di questa scelta? È che sicuramente queste serie avranno risultato, finito o infinito. Infatti la successione delle somme parziali risulta crescente, e si dimostra facilmente:

$$s_n < s_{n+1} \quad s_{n+1} - s_n > 0 \quad a_{n+1} > 0$$

Il che è vero per ipotesi. Ripercorrendo le uguaglianze al contrario si ha la dimostrazione cercata. Quindi, per il Teorema sull'esistenza del limite delle successioni monotòne (4.1.3), abbiamo che la serie o converge ad un valore finito, oppure a più infinito. Ovviamente tutti i discorsi che si faranno in seguito varranno anche per serie a tutti termini negativi.

Inoltre, sussiste anche il seguente criterio, chiamato:

Criterio 4.2.2 (Criterio d'equivalenza (o criterio del confronto asintotico)) *Se $a_n \sim b_n$, allora le serie*

$$\sum_{n=0}^{\infty} a_n \quad e \quad \sum_{n=0}^{\infty} b_n$$

hanno lo stesso comportamento, ovvero sono entrambe o divergenti o convergenti

Dimostreremo ciò fra poco.

Esempio 4.2.2 *Dimostrare che la serie*

$$\sum_{n=1}^{\infty} \frac{1}{n^2}$$

è convergente.

Per dimostrare ciò prima dimostreremo che un'altra serie è convergente, per la precisione la cosiddetta serie di Mengoli. *Questa è la serie:*

$$\sum_{n=1}^{\infty} \frac{1}{n(n+1)}$$

Osserviamo inanzitutto che ciascun termine si può riscrivere come:

$$\frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1}$$

Quindi, troviamo la somma parziale n-sima:

$$s_n = \frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \frac{1}{3 \cdot 4} + \cdots + \frac{1}{n(n+1)}$$

che si può anche scrivere come:

$$s_n = \underbrace{1 - \frac{1}{2}}_{\frac{1}{1\cdot 2}} + \underbrace{\frac{1}{2} - \frac{1}{3}}_{\frac{1}{2\cdot 3}} + \underbrace{\frac{1}{3} - \frac{1}{4}}_{\frac{1}{3\cdot 4}} + \cdots + \underbrace{\frac{1}{n} - \frac{1}{n+1}}_{\frac{1}{n(n+1)}} = 1 - \frac{1}{n+1}$$

A questo punto possiamo calcolare il limite di s_n per n che tende ad infinito, e otteniamo:

$$\lim_{n\to\infty} s_n = 1$$

Il quale quindi è anche il risultato della serie.

Torniamo ora alla serie originale, e verifichiamo che i termini della sua successioni sono asintotici con quelli della serie di Mengoli:

$$\frac{\frac{1}{n^2}}{\frac{1}{n(n+1)}} = \frac{1}{n^2} \cdot (n^2 + n) = \frac{n^2 + n}{n^2}$$

che è una successione che chiaramente tende ad 1. Le due serie sono quindi convergenti e, per il criterio di equivalenza, hanno lo stesso comportamento: la serie armonica data quindi converge.

Facciamo notare inoltre che solamente il comportamente è lo stesso, ma non il valore a cui tendono: la serie di Mengoli tende ad 1, quella armonica considerata al ben più strano valore $\pi^2/6$.

Il criterio di equivalenza ci consente di determinare molto spesso il carattere di una serie senza particolari difficoltà, anche quando la serie in questione appare piuttosto complessa: basta riuscire a trovarne un'altra asintotica alla prima.

Esempio 4.2.3 *Determinare che la serie*

$$\sum_{n=0}^{\infty} \frac{n^2+1}{n^4+\sqrt{n+2}}$$

converge o meno.

Osserviamo senza molti problemi che la successione dei termini della serie è asintotica a $1/n^2$ *– dato che abbiamo dimostrato la convergenza di questa serie, anche quella data risulta convergente.*

Il criterio di equivalenza è un caso particolare del più vasto principio seguente:

Criterio 4.2.3 (Criterio del confronto) *Date due serie*

$$\sum_{n=0}^{\infty} a_n \quad e \quad \sum_{n=0}^{\infty} b_n$$

a termini positivi, se $\exists c > 0 : a_n \leq c \cdot b_n, \forall n \in \mathbb{N}$*, allora abbiamo che:*

- *Se* $\sum b_n$ *converge, allora* $\sum a_n$ *converge.*
- *Se* $\sum a_n$ *diverge, allora* $\sum b_n$ *diverge.*

Dimostrare il criterio di equivalenza a partire da questo è abbastanza facile.

Inanzitutto, abbiamo che, se $a_n \sim b_n$, allora:

$$\lim_{n \to +\infty} \frac{a_n}{b_n} = 1$$

Il che, utilizzando la definizione di limite, significa che

$$\forall \varepsilon \geq 0, 1 - \varepsilon < \frac{a_n}{b_n} < 1 + \varepsilon, n > n_\varepsilon$$

Il che, se vale per qualunque ε, varrà in particolare per qualunque ε che renda la differenza $1 - \varepsilon$ positiva, ovvero per qualunque $\varepsilon \leq 1$.

Ora, i rapporti di successione non considerati sono quelli con coefficiente $\leq n_\varepsilon$, quindi comunque un numero finito di valori. Quindi, potrò sempre avere tra di loro un minimo m ed un massimo M:

$$0 < m \leq \frac{a_n}{b_n} \leq M, n \leq n_\varepsilon$$

Ora, considerando entrambe le disuguaglianze scritte, si trova che in generale i rapporti a_n/b_n saranno maggiore del più piccolo tra $1 - \varepsilon > 0$ ed m, e saranno minori del più grande tra $1 + \varepsilon$ ed M, quindi abbiamo che:

$$0 < min(1 - \varepsilon, m) \leq \frac{a_n}{b_n} \leq max(M, 1 + \varepsilon)$$

Ora, chiamiamo il valore minimo k ed il massimo h. Abbiamo quindi:

$$0 < k \leq \frac{a_n}{b_n} \leq h$$

Ora, il "0 <" ad inizio disuguaglianza ci assicura che lavoriamo con valori positivi e non dobbiamo quindi neanche preoccuparci di cambiamenti di verso della disuguaglianza con divisioni e moltiplicazioni.

Dal primo termine ricaviamo quindi che

$$a_n \leq h b_n, \forall n \in \mathbb{N}$$

e quindi per il criterio del confronto ($h > 0$), questo significa che:

- Convergenza di $b_n \implies$ Convergenza a_n
- Divergenza di $a_n \implies$ Divergenza b_n

Dalla prima uguaglianza troviamo invece che:

$$b_n \leq \frac{1}{k} a_n$$

E quindi, sempre per il criterio del confronto ($1/k > 0$), abbiamo che:

- Convergenza di $a_n \implies$ Convergenza b_n
- Divergenza di $b_n \implies$ Divergenza a_n

Ora, unendo i due blocchi di affermazioni abbiamo:

- Convergenza di $a_n \iff$ Convergenza b_n
- Divergenza di $a_n \iff$ Divergenza b_n

E quindi, in pratica, le due serie hanno lo stesso comportamento. CVD.

4.2.3 Criteri di convergenza

Esistono anche alcuni altri criteri utili per la determinazione della convergenza delle serie:

Criterio 4.2.4 (Criterio del rapporto (per le serie a termini positivi)) *Data una serie $\sum a_n$ a termini positivi, allora abbiamo che, se l esiste:*

$$\lim_{n\to+\infty} \frac{a_{n+1}}{a_n} = \begin{cases} 0 \le l < 1 \Rightarrow \sum a_n \text{ converge} \\ l > 1, l = +\infty \Rightarrow \sum a_n \text{ diverge} \\ l = 1 \Rightarrow \text{nulla si può dire sul risultato} \end{cases}$$

Come per il corrispondente criterio per le successioni, anche in questo caso abbiamo che, nel caso di $l = 1$, nulla è possibile dire sul risultato: lo mostrano chiaramente due semplici serie come $1/n$ e $1/n^2$, la prima divergente e la seconda convergente come già sappiamo, che però hanno entrambe $l = 1$.

Esempio 4.2.4 *Determinare il carattere della serie*

$$\sum_{n=1}^{\infty} \frac{n!}{n^n}$$

Utilizzando il criterio del rapporto:

$$\frac{a_{n+1}}{a_n} = \frac{\frac{(n+1)!}{(n+1)^{n+1}}}{\frac{n!}{n^n}} = \frac{(n+1)n!}{(n+1)(n+1)^n} \cdot \frac{n^n}{n!} = \left(\frac{n}{n+1}\right)^n = \frac{1}{\left(1+\frac{1}{n}\right)^n} \to \frac{1}{e}$$

Dato che $0 < 1/e < 1$, la serie converge.

Altro criterio utile è il seguente:

Criterio 4.2.5 (Criterio della radice (per le serie)) *Sia $\sum a_n$ una serie a termini positivi, allora abbiamo che, se l esiste:*

$$\lim_{n\to+\infty} \sqrt[n]{a_n} = l = \begin{cases} l < 1 \Rightarrow \sum a_n \text{ converge} \\ l > 1, l = +\infty \Rightarrow \sum a_n \text{ diverge} \\ l = 1 \Rightarrow \text{nulla si può dire sul risultato} \end{cases}$$

4.2.4 Serie a termini di segno variabile

Finora abbiamo studiato serie a termini di segno costante. Se prendiamo invece serie con segno variabile, le cose cambiano. Intanto però si può sempre dire che:

Criterio 4.2.6 (Criterio della convergenza assoluta) *Se una serie è* assolutamente convergente*, ovvero se converge la serie che ha per termini i valori assoluti dei termini della serie considerata, allora la serie è anche* semplicemente convergente*, ovvero converge essa stessa.*

$$\text{Converge } \sum_{n=0}^{\infty} |a_n| \quad \Longrightarrow \quad \sum_{n=0}^{\infty} a_n$$

Questo teorema però non ci dice nulla nel caso in cui la serie sia assolutamente *divergente* – questo significa che in quella situazione la serie stessa potrebbe essere sia convergente che divergente.

Esempio 4.2.5 *Studiare la serie*

$$\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}$$

attraverso il criterio della convergenza assoluta.

Questa serie è a termini positivi e negativi, per la precisione è una serie a termini alternati*, ovvero alternativamente positivi e negativi.*

Se analizziamo la serie dei valori assoluti, troviamo:

$$\sum_{n=1}^{\infty} \left|\frac{(-1)^{n-1}}{n}\right| = \sum_{n=1}^{\infty} \frac{|(-1)^{n-1}|}{n} = \sum_{n=1}^{\infty} \frac{1}{n}$$

che noi sappiamo divergente. Quindi, per mezzo di questo criterio, in realtà noi non possiamo dire nulla sulla serie di partenza*. Anzi, questo particolare esempio porta come esempio una serie che è* convergente*, ed è facile convincersene intuitivamente esaminando i suoi termini:*

$$1, -\frac{1}{2}, \frac{1}{3}, -\frac{1}{4}, \dots$$

Si può notare che la somma dei primi due termini è senz'altro minore del primo termine, quella dei primi tre termini è maggiore della somma dei primi due, quella dei primi quattro è minore della somma dei primi tre, e così via: quindi abbiamo trovato che la successione delle somme parziali è limitata, ed è facile mostrare che ammette anche limite, che deve essere quindi finito. Più avanti scopriremo che questo limite è proprio $\ln 2$*.*

Esempio 4.2.6 *Determinare il carattere della successione*

$$\sum_{n=1}^{\infty} \frac{(-1)^{p(n)}}{n^2}$$

dove $p(n)$ è così definita:

$$p(n) = \begin{cases} 1 \text{ se } n \text{ è primo} \\ 0 \text{ se } n \text{ non è primo} \end{cases}$$

Questa serie, al contrario della precedente, non è a termini alternati, sebbene rimanga una serie a termini positivi e negativi. Anche se il suo studio sembra di notevole complessità (ed in effetti il suo calcolo lo sarebbe), se ci si accontenta di sapere se essa è convergente o divergente basta veramente poco – esaminando infatti la serie dei valori assoluti abbiamo:

$$\sum_{n=1}^{\infty} \left| \frac{(-1)^{p(n)}}{n^2} \right| = \sum_{n=1}^{\infty} \frac{|(-1)^{p(n)}|}{n^2} = \sum_{n=1}^{\infty} \frac{1}{n^2}$$

che è una serie convergente. Quindi, dato che la serie è assolutamente convergente, per il criterio poco fa esposto, è anche semplicemente convergente.

Se ci troviamo di fronte ad una successione che non solo è, genericamente, non una serie a termini positivi, ma è a termini alternati, ovvero, come detto prima, a termini alternativamente positivi e negativi, allora vale il seguente criterio:

Criterio 4.2.7 (Criterio di Leibniz) *Se per la serie a termini alterni*

$$\sum_{n=0}^{\infty} a_n$$

valgono le seguenti proposizioni:

- $\sum |a_n|$ *è decrescente*
- $|a_n|$ *tende a 0*

che possono anche essere scritte assieme con la notazione:

$$a_n \searrow 0$$

allora la serie è convergente.

Come si nota questo criterio ha richieste meno restrittive rispetto a quello della convergenza assoluta, però vale solo per le serie a termini alternati: non si sarebbe potuto applicare ad esempio alla serie esaminata prima: $\sum a_n^{p(n)}$.

Esempio 4.2.7 *Determinare il carattere della serie*

$$\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}$$

Con il criterio della convergenza assoluta non eravamo riusciti a determinare il carattere di questa serie, vediamo ora a cosa riusciamo ad arrivare. Prima di tutto calcoliamo $|a_n|$:

$$|a_n| = \left| \frac{(-1)^{n-1}}{n} \right| = \frac{1}{n}$$

(n è sempre positivo). A questo punto, vediamo che le due ipotesi del criterio di Leibniz sono soddisfatte, infatti $1/n$ è decrescente e tende a 0, quindi la serie presa in esame converge.

Piccolo riassunto dei criteri di convergenza delle serie:

- Se una serie è convergente, il termine generico a_n deve tendere a 0, ovvero:

$$\text{Convergenza di } \sum a_n \quad \Longrightarrow \quad a_n \to 0$$

- Se una serie converge assolutamente, allora converge anche semplicemente, ovvero:

$$\text{Convergenza di } \sum |a_n| \quad \Longrightarrow \quad \text{Convergenza di } \sum a_n$$

(Criterio della convergenza assoluta - 4.2.6)

- Se abbiamo una serie a termini positivi, allora valgono i criteri del confronto (4.2.3), di equivalenza (4.2.2), del rapporto (4.2.4) e della radice (4.2.5).
- Se abbiamo una serie a termini alterni, vale il criterio di Leibniz (4.2.7).

In generale, quindi, il percorso da seguire per determinare la convergenza o meno di una serie, è il seguente:

- Verificare che $a_n \to 0$, se così non è, la serie è divergente.
- Controllare se la serie è a termini tutti positivi (o negativi): in questo caso si possono utilizzare i criteri per queste serie – tipicamente quelli del rapporto/radice e di equivalenza.
- Nel caso abbia segno non costante, controllare se è a termini di segno alternato: in questo caso si può usare il criterio di Leibniz
- Altrimenti, l'unica via disponibile è quella di esaminare se la serie è assolutamente convergente, e quindi ci si riduce a studiare la serie a termini positivi $\sum |a_n|$, e in questo caso si sa che la serie è anche semplicemente convergente – in caso contrario, non si può dire nulla.

Esempio 4.2.8 *Determinare per quali $x \in \mathbb{R}$ è convergente la serie*

$$\sum_{n=1}^{\infty} \frac{x^n}{n}$$

Cominciamo con l'osservare che per $x > 0$, abbiamo una serie a termini tutti positivi, per $x < 0$ invece una serie a segno alternato. Per il caso specifico di $x = 0$ abbiamo una serie particolarmente semplice:

$$\sum_{n=1}^{\infty} \frac{0^n}{n} = \sum_{n=1}^{\infty} 0 = 0$$

e che quindi è convergente.

Per esaminare il caso $x > 0$, possiamo utilizzare uno dei vari criteri che abbiamo a nostra disposizione per le serie a termini positivi – usiamo in questo caso ad esempio quello del rapporto (anche con quello della radice si arriverebbe a risultati analoghi):

$$\frac{a_{n+1}}{a_n} = \frac{\frac{x^{n+1}}{n+1}}{\frac{x^n}{n}} = \frac{x \cdot x^n}{n+1} \cdot \frac{n}{x^n} = \frac{n}{n+1} \cdot x \to x$$

(questo perché $n/(n+1) \to 1$). Quindi, applicando il criterio del rapporto, otteniamo che per $0 \leq x < 1$ la serie converge (per 0 avevamo già d'altronde avuto conferma), per $x > 1$ diverge, ed infine per $x = 1$, nulla si può dire. Per le x positive, quindi, ci rimane ancora da analizzare il caso $x = 1$, ma questo non è altro che:

$$\sum_{n=1}^{\infty} \frac{1^n}{n} = \sum_{n=1}^{\infty} \frac{1}{n}$$

che è la serie armonica semplice, una serie divergente. Quindi, per le x positive, abbiamo intanto che la serie è convergente solo in [0;1).

Guardiamo ora le x negative: abbiamo detto che è una serie a termini alternati, per iniziare possiamo intanto esaminare la convergenza assoluta:

$$|a_n| = \left| \frac{x^n}{n} \right| = \frac{|x|^n}{n}$$

Ma questa non è altro che la serie che abbiamo già esaminato, e possiamo riutilizzare gli stessi risultati già ottenuti, che ci dicono che è convergente per $x \in$[0;1) e divergente per le $x \geq 1$. Quindi, tornando alla nostra serie di partenza, possiamo dire che essa è convergente per $x \in$(-1;0], mentre per le altre x, nulla si può dire. *Ricordiamo infatti che il criterio della convergenza assoluta ci dice solo che una serie assolutamente convergente è anche semplicemente convergente, ma non ci dice nulla sul comportamente semplice della serie se il suo comportamente assoluto è di divergenza. Dobbiamo quindi trovare un altro sistema per esaminare questa serie: prendiamo inanzitutto il caso limite $x = -1$, che possiamo esaminare con il criterio di Leibniz dato che, come detto, è a segno alternato:*

$$|a_n| = \left| \frac{(-1)^n}{n} \right| = \frac{1}{n}$$

che è una serie decrescente che tende a 0: quindi per il caso $x = -1$ la serie è ancora convergente.

Per le $x < -1$ osserviamo molto semplicemente che non è rispettata la condizione necessaria di convergenza, ovvero $a_n \to 0$. Infatti, dato che x^n è un infinito di ordine superiore di n per $x > 1$, il loro rapporo tende ad infinito nel caso $x > 1$ - noi abbiamo qui $x < -1$, per il quale la potenza assume gli stessi valori del caso $x > 1$ ma con segni alternati: quindi il termine generico a_n per $n < -1$ assume valori sempre più grandi alternativamente positivi e negativi, e non ha quindi limite. La serie è quindi divergente in questo caso.

Riassumendo, abbiamo che la risposta al problema (ovvero l'intervallo a cui deve appartenere x per avere una serie convergente) è:

$$-1 \leq x < 1$$

o anche, che è la stessa cosa:

$$x \in [-1; 1)$$

Esempio 4.2.9 *Determinare il carattere della serie*

$$\sum_{n=1}^{\infty} \frac{\cos\left(\frac{\pi}{2}n\right)}{n}$$

Il primo passo da fare è tentare di semplificare l'espressione che esprime questa successione: osserviamo infatti che il numeratore di a_n assume ciclicamente i valori 0, -1, 0 e +1. Quindi avremo che i termini di posto dispari non compaiono nella somma finale della serie, e possono quindi essere ignorati, e rimangono solo quelli pari con segno alternatativamente negativo e positivo; possiamo quindi riscrivere la serie come:

$$\sum_{k=1}^{\infty} \frac{(-1)^k}{2k}$$

(l'uso di un indice di nome diverso è dovuto al fatto che $a_n \neq a_k$ quando $n = k$, dato che k percorre il "doppio di strada" rispetto ad n).

Ora possiamo utilizzare il criterio di Leibniz per studiare la serie:

$$|a_n| = \frac{1}{2k}$$

che è una serie decrescente e che tende a 0 - la serie data quindi converge.

Esempio 4.2.10 *Studiare il carattere della serie*

$$\sum_{n=1}^{\infty} \frac{\sin n + (-1)^n n}{n^2}$$

A prima vista questa serie appare decisamente complicata – infatti i valori che $\sin n$ assume al variare di n sono sempre diversi, e l'unica cosa che li accomuna è che, ovviamente, sono compresi tra -1 ed 1 – ma questo è l'unico dato di cui avremo bisogno, come vedremo.

Possiamo seguire questa strada: se riusciamo a dimostrare che le due serie:

$$\sum_{n=1}^{\infty} \frac{\sin n}{n^2}$$

e

$$\sum_{n=1}^{\infty} \frac{(-1)^n n}{n^2}$$

convergono, convergerà ovviamente anche la loro somma, che è per l'appunto la serie data da studiare. Partiamo quindi con l'esaminare la prima delle due serie, i cui termini chiameremo b_n. È una serie non a termini positivi, e nemmeno alternati: l'unico criterio che possiamo usare quindi è quello della convergenza assoluta, ovvero studiare la serie:

$$\sum_{n=1}^{\infty} \frac{|\sin n|}{n^2}$$

Osserviamo quindi che $|\sin n|$ è sempre minore di uno, e quindi possiamo dire che:

$$\frac{|\sin n|}{n^2} \leq \frac{1}{n^2}$$

e quindi, per il criterio del confronto, dato che la serie

$$\sum_{n=1}^{\infty} \frac{1}{n^2}$$

converge, converge anche l'altra serie. Quindi, dato che la serie $\sum b_n$ è assolutamente convergente, allora è anche semplicemente convergente.

Rimane ora da dimostrare che è convergente pure l'altra serie. Ma questa, semplificata, risulta essere uguale a:

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n}$$

che abbiamo già dimostrato in altri esempi col criterio di Leibniz che è convergente.

Quindi, dato che le due serie prese in considerazione sono convergente, è convergente anche la loro somma. La serie data è quindi convergente.

Capitolo 5

Limite di funzione reale

5.1 Definizione di limite di funzione reale

Dopo aver definito, nel precendente capitolo, il limite di una successione, passiamo ora a definire il limite di una funzione reale, utilizzando la gli strumenti già a nostra disposizione.

Si dice *funzione reale* una funzione che ha per dominio un sottoinsieme (proprio od improprio) di $\mathbb{R}$ e per codominio proprio $\mathbb{R}$.

La definizione "intuitiva" di limite anche in questo caso è facile: il limite, per x che tende ad un certo valore x_0, sia esso finito o infinito, della funzione $f(x)$, è quel valore a cui la funzione $f(x)$ si avvicina quando x si avvicina ad x_0.

Per poter formulare una definizione matematica di questi concetti, partiamo col definire cosa significa che *il punto c è raggiungibile da A (un insieme di numeri)* o anche, che è la stessa cosa, che *c è un punto di accumulazione di A.*

c è raggiungibile da A se e solo se esiste almeno una successione x_n tale che:

- $x_n \in A, \forall n \in \mathbb{N}$
- $x_n \neq c, \forall n \in \mathbb{N}$
- $x_n \to c$

Esempio 5.1.1 *Dato*

$$A = (-1; 1] + \{3\}$$

dire se i seguenti punti sono o meno raggiungibili da A:

1. $c = 0$
2. $c = 1$
3. $c = -1$
4. $c = 3$

Inanzitutto, osserviamo A: questo insieme è costituito da tutti i punti compresi tra -1 *(escluso) ed* 1 *(incluso) con l'aggiunta del punto* 3.

Per quanto riguarda il punto 0*, se prendiamo la successione* $x_n = e^{-n}$*, abbiamo una successione che appartiene per ogni* n *naturale ad A, che non è mai uguale a* 0*, ma che tende a* 0*.* 0 *è quindi raggiungibile da A: possiamo generalizzare dicendo che un punto interno ad un intervallo è sempre raggiungibile da quell'intervallo.*

Riguardo al punto 1*, anche questo punto appartiene all'intervallo, ma non è propriamente interno ad esso, in quanto è sul suo confine, tuttavia anche per questo punto si può trovare una successione* x_n *adatta – ad esempio,* $1 - e^{-n}$ *(la verifica è semplice). Quindi, anche* 1 *è raggiungibile da A.*

Ora prendiamo il punto -1*: questo punto* non *appartiene ad A, eppure è comunque raggiungibile da esso, infatti se prendiamo, ad esempio, la successione* $e^{-n} - 1$ *ha tutti termini appartenenti ad A, mai uguali a* -1 *e che però tendono ad esso. Quindi, dai tre risultati precedenti, possiamo generalizzare dicendo che* i punti appartenenti ad un intervallo ed i suoi estremi, che siano inclusi o esclusi, sono raggiungibile dall'intervallo stesso.

Infine, prendiamo il punto 3*: questo punto, pur appartenendo all'insieme A non è raggiungibile da esso, infatti non è possibile costruire nessuna successione che tenda a* 3 *ed allo stesso tempo abbia tutti i suoi termini che appartengono ad A – questo poiché, per tendere a* 3*, dovrebbero i suoi termini avvicinarsi indefinitamente a questo valore, ma se lo facessero dovrebbero senz'altro uscire da A.*

Abbiamo quindi visto che l'appartenenza o meno all'insieme A non condiziona la raggiungibilità *dallo stesso.*

A questo punto, possiamo dare la definizione di limite.

Si dice che il limite di $f(x)$ per x che tende a c è l (dove sia c che l possono essere valori sia finiti che infiniti e c è un punto raggiungibile dal dominio di $f(x)$), e si scrive:

$$\lim_{x \to c} f(x) = l$$

se e solo se:

$$\forall (x_n)_{n \in \mathbb{N}} : f(x_n) \to l$$

Dove x_n è una successione che rispetta le stesse condizioni date nella definizione di raggiungibilità, ovvero:

- $x_n \in A, \forall n \in \mathbb{N}$
- $x_n \neq c, \forall n \in \mathbb{N}$
- $x_n \to c$

Dove A è il dominio di $f(x)$.

Da notare che f è una funzione, ma $f(x_n)$ è una successione: in effetti x_n è una successione ed f non fa altro che cambiare i risultati della successione.

Esempio 5.1.2 *Dimostrare che*

$$\lim_{x \to 0} \frac{1}{x^2} = +\infty$$

Studiamo inanzitutto il dominio della nostra $f(x)$, che risulta essere

$$\mathbb{R} - \{0\}$$

Quindi il punto c, che è 0, non appartiene al dominio - ma è raggiungibile da esso? Basta trovare una successione che rispetti le tre condizioni poste, e ciò è molto facile: alcune delle successioni che soddisfano ciò sono ad esempio $1/n$, $1/n^2$, $(-1)^n/n^2$, ...

Ora, dobbiamo dimostrare che, qualunque sia la nostra x_n, la successione $f(x_n)$ tende a 0, ovvero che:

$$\frac{1}{x_n^2} \to +\infty$$

Ma questo è molto facile: utilizzando infatti le regole di calcolo che conosciamo, sappiamo per ipotesi che $x_n \to 0$ (guarda la definizione di limite!), quindi il suo quadrato tenderà a zero pure lui, e per di più sempre da termini di valore positivo, quindi il suo reciproco tenderà a più infinito. Ecco dimostrato il limite in questione.

Esempio 5.1.3 *Dimostrare che*

$$\lim_{x \to \pm\infty} \frac{1}{x^2} = 0$$

I punti più o meno infinito sono senz'altro raggiungibili da A, infatti anche in questo caso sono svariate le successioni che tendono a tale valore, come ad esempio n, n^2, 2^n, ..., e con il segno invertito per arrivare a meno infinito.

Consideriamo il caso di $c = +\infty$ – il caso di $c = -\infty$ è del tutto analogo.

Presa una qualunque successione x_n che soddisfi le condizioni date, abbiamo che sicuramente $x_n \to +\infty$, quindi $x_n^2 \to +\infty$, $1/x_n^2 \to 0$, e quindi che

$$f(x_n) \to 0$$

Ecco anche questa volta dimostrato il limite.

Esempio 5.1.4 *Determinare il valore di*

$$\lim_{x \to 0} \frac{1}{x}$$

Il dominio che abbiamo è $\mathbb{R} - \{0\}$, tuttavia da questo insieme, come abbiamo già mostrato, il punto 0 è raggiungibile, e quindi accettabile nella definizione di limite.

Prendiamo come al solito una qualunque successione $x_n \to 0, \neq 0$: il suo inverso, $1/x_n$, è una forma di indeterminazione $(1/0)$, di cui quindi non siamo in grado di determinare il limite. Ebbene, il limite della funzione presa in esame, in realtà non esiste: se infatti prendo per $x_n = 1/n$ trovo come limite $+\infty$, se prendo invece $x_n = -1/n$ trovo $-\infty$, se infine prendo $(-1)^n/n$, addirittura non ho limite. Come si può notare sicuramente non tutte le successioni che si possono prendere danno come risultato un unico limite – quindi, il limite non esiste. (è facile capire da questo esempio che la proprietà di unicità del limite enunciata per le serie vale anche per le funzioni).

Esempio 5.1.5 *Determinare*

$$\lim_{x\to 0} f(x)$$

Dove

$$f(x) = \begin{cases} 1+x & x \geq 0 \\ -x & x < 0 \end{cases}$$

È facile dimostrare che questo limite non esiste. Infatti, se prendo una $a_n > 0$, ho che $f(a_n) = 1 + a_n$ che tende chiaramente ad 1; se prendo invece una $b_n < 0$, ho $f(b_n) = -b_n$ che invece tende a 0. Quindi, non tutte le successioni date tendono ad un unico valore, e quindi il limite non esiste.

Quest'ultimo esempio presenta un caso in cui il limite esisterebbe se si prendessero solo valori maggiori di c o minori di tale valore (ad esempio nel caso precendente, il limite a destra di c è $+1$, mentre a sinistra è 0). Esiste un modo per indicare ciò: ad esempio, se voglio parlare del limite considerando la funzione solo a destra del punto c, allora utilizzerò il limite *destro*, che si indica così:

$$\lim_{x\to c^+} f(x)$$

E che, come unica differenza nella definizione, ha, nell'elenco delle proprietà a cui deve sottostare x_n, il fatto che essa non deve essere solo diversa da c per ogni suo termine, ma anche *maggiore* di esso:

- $x_n \in A$
- $x_n > c$
- $x_n \to c$

La definizione di limite *sinistro* è del tutto analoga.

È facile dimostrare che

$$\lim_{x\to c} f(x) = \lim_{x\to c^+} f(x) = \lim_{x\to c^-} f(x)$$

E che tale limite esiste solo se esistono uguali i due limiti destro e sinistro.

5.2 Funzioni continue

Una funzione $f(x)$ viene detta *continua* nel punto $c \in A$ (dominio della funzione) e raggiungibile da A stesso, se e solo se $\lim_{x\to c} f(x) = f(c)$. Una funzione $f(c)$ viene poi detta *continua su A* se è continua in ogni punto di tale insieme. Una funzione continua sul suo dominio viene detta semplicemente *continua*.

La continuità in un punto, come è facile convincersi, non dipende solo dal comportamente della funzione in quel punto, ma proprio da come quella funzione è strutturata intorno a quel punto. Se prendiamo ad esempio la funzione

$$f(x) = \frac{1}{x^2}$$

si trova facilmente che non è continua nel punto 0. Anche se estendiamo la definizione di $f(x)$ in modo che sia definita anche in 0, ad esempio così :

$$f(x) = \begin{cases} \dfrac{1}{x^2}, & x \neq 0 \\ \\ 0, & x = 0 \end{cases}$$

Si nota che la situazione non cambia, perchè non si riesce a trovare il limite della funzione nel punto 0.

Tutte le funzioni elementari sono continue, ovvero continue sul loro dominio; tali funzioni sono:

$$x^k, \alpha^x, \ln_\alpha x, \sin x, \cos x$$

Con le solite restrizioni per i valori i k ed α. D'altronde, quasi ogni operazione di composizione tra funzioni continue crea a sua volta funzioni continue – detto in altre parole, la continuità è una proprietà stabile rispetto a molti operatori.

Per la precisione, se f e g sono due funzioni continue, allora:

1. $f + g$ è continua
2. $f - g$ è continua
3. λf è continua (λ è un qualunque valore reale)
4. $f \cdot g$ è continua
5. $1/f$ è continua se $f \neq 0$

6. $f \circ g$ (o anche $f(g)$) è continua, se $f(A) \subseteq C$, dove A è il dominio di f (quindi $f(A)$ è l'immagine del dominio di f) e C il dominio di g.

7. f^{-1} (la funzione inversa di f) è continua, se f è una funzione definita da un intervallo ad un intervallo ed è biunivoca.

Riguardo alla composizione di funzioni: date due funzioni f e g, condizione necessaria e sufficiente perchè sia possibile fare la composizione di f con g, indicata da $f \circ g$ o da $f(g)$, è che $f(A) \subseteq C$ (la terminologia usata è la stessa di quella utilizzata poco sopra). In genere, l'immagine del dominio di una funzione è sottoinsieme del suo codominio.

L'operatore di composizione non è commutativo: ad esempio, $\sin x^2 \neq \sin^2 x$.

Per quanto riguarda l'invertibilità, abbiamo detto che la condizione necessaria e sufficiente è che la funziona sia biunivoca, ovvero che sia una mappa uno-ad-uno tra il dominio e il codominio: ad ogni elemento del dominio corrisponde un elemento del codominio e vicevera. Un altro modo per esprimere questo concetto è dire che la funzione in questione è sia iniettiva che suriettiva. La definizione di funzione iniettiva

$$f : A \to B$$

è:

$$\forall x_1, x_2 \in A \mid x_1 \neq x_2 \;\Rightarrow\; f(x_1) \neq f(x_2)$$

Ovvero, presi due punti del dominio, essi non possono avere la stessa immagine. Questa è la condizione più importante da verificare per assicurarsi l'invertibilità. Inoltre la suriettività di una funzione dice questo:

$$\forall y \in B \mid \exists x \in A \mid f(x) = y$$

Ovvero, in termini più semplici, tutti i punti del codominio sono immagini di punti del dominio, e quindi $f(A) \equiv B$. Con queste due condizioni ci si assicura che la funzione inversa sia effettivamente una funzione, in quanto ogni elemento del codominio non ha più di una immagine (iniettività della funzione di partenza) nè meno di una immagine (suriettività della funzione di partenza). In questo modo anche la funzione inversa è biunivoca.

Un altro paio di fatti utili sono i seguenti: se una funzione è monotona strettamente crescente (o strettamente descrescente) allora è invertibile (dall'immagine del suo dominio al dominio). Inoltre una funzione continua, come già detto, se ammette funzione inversa, essa è ancora continua.

Esempio 5.2.1 *Dimostrare che*

$$\sqrt{\frac{x^2-1}{x^2+1}}$$

è una funzione continua (su tutto il suo dominio)

Questo esempio mostra l'utilizzo di varie regole di composizione di funzioni continue. Inanzitutto, il dominio della nostra $f(x)$ è:

$$\begin{cases} x^2+1 \neq 0 \\ \dfrac{x^2-1}{x^2+1} \geq 0 \end{cases} \quad \Longrightarrow \quad x^2-1 \geq 1 \quad \Longrightarrow \quad x \in (-\infty,-1] \cup [1,+\infty)$$

Cominciamo ad esaminare la continuità della funzione ora, e verifichiamo che è tale sul suo dominio: x^2 è una potenza, quindi una funzione continua - la sua somma o differenza con altre funzioni (costanti, in questo caso, poiché sono $+1$ e -1) daranno come risultato quindi funzioni continue (regole 1 e 2): x^2-1 ed x^2+1 sono quindi continue. In generale, è facile da capire che tutti i polinomi sono funzioni continue.

Inoltre, dato che x^2+1 è sempre positivo, allora il rapporto

$$\frac{x^2-1}{x^2+1}$$

(per le regole 4 e 5) è una funzione continua. Ora dobbiamo dimostrare che la radice di tale funzione è funzione continua.

Adesso, per determinare se la sua radice produce una funzione continua, dobbiamo verificare se le composizione della funzione "radice quadrata" e della funzione appena trovata è possibile, e anche se la funzione "radice quadrata" è continua. Per il primo punto, occorre trovare $f(A)$, ovvero l'immagine del dominio, che è $[0;+\infty)$. Questo insieme è esattamente il dominio della radice quadrata, e quindi la loro composizione, se la radice risulta essere continua, è una funzione continua (regola 6).

La radice è la funzione inversa dell'elevamento al quadrato: L'elevamento al quadrato è una funzione biunivoca, se preso sull'intervallo $[0;+\infty)$, e quindi su questo intervallo è invertibile, dandoci così la funzione radice quadrato, che per la regola 7 è quindi a sua volta continua. CVD.

5.3 Teoremi sulle funzioni continue

Esistono tre fondamentali teoremi sulle funzioni continue, che verranno presentati in questa sezione.

Teorema 5.3.1 (Teorema degli zeri) *Data $f : [a, b] \to \mathbb{R}$, con f continua e discorde agli estremi dell'intervallo ($f(a)f(b) < 0$), allora $\exists c \in (a, b) : f(c) = 0$.*

In pratica questo teorema assicura che, alle condizioni date, il grafico della funzione interseca sempre almeno una volta l'asse dell'ascisse, e quindi l'equazione $f(c) = 0$ ha almeno una soluzione.

La dimostrazione che si da di solito di questo teorema è *costruttiva*, ovvero non solo dimostra la veridicità del teorema, ma indica anche un modo (un "algoritmo") per arrivare al calcolo del risultato. Nel caso particolare, quel che la dimostrazione fa è individuare un metodo sempre funzionante per calcolare il nostro valore di c.

Per questa dimostrazione prendiamo il caso di $f(a) > 0$ ed $f(b) < 0$, per il caso opposto la dimostrazione sarà ovviamente strutturata allo stesso modo.

a questo punto calcoliamo $c = \frac{a+b}{2}$: possiamo avere due casi:

1. $f(c) = 0$, ed in questo caso abbiamo trovato il valore che cercavamo
2. $f(c) < 0$ o $f(c) > 0$

in questo caso, al posto di esaminare l'intervallo $[a, b]$, esaminiamo l'intervallo $[a, \frac{a+b}{2}]$ (nel primo caso) oppure $[\frac{a+b}{2}, b]$ (nel secondo caso). Se ora chiamiamo di nuovo a l'estremo inferiore di tale intervallo e b l'estremo superiore, ci ritroviamo di nuovo nella situazione di partenza, ma con un intervallo più piccolo, per la precisione grande esattamente la metà.

Ci possiamo trovare in due condizioni ovviamente: o in un numero finito di passi ci troviamo nel caso 1 tra quelli elencati precedentemente, o non vi arriviamo mai. Nel primo caso, come già detto, abbiamo la nostra soluzione, mentre nel caso contrario la possiamo trovare in un altro modo.

Quel che facciamo, continuando indefinitamente a costruire intervalli sempre più piccoli, è di costruire una successione di estremi inferiori e superiori dell'intervallo nel quale cerchiamo la soluzione; chiamiamo la successione degli estremi inferiori

$$\{a_n\}_{n\in\mathbb{N}}$$

e quella degli estremi superiori

$$\{b_n\}_{n\in\mathbb{N}}$$

Ovviamente

$$a_1 = a; \quad b_1 = b$$

Poichè sono i valori con cui partiamo. La lunghezza dell'intervallo così definito è a sua volta una successione, molto più semplice di queste prime due; al primo passaggio ha lunghezza pari a:

$$i_1 = a - b$$

e per ogni passaggio successivo, come già detto, dimezza di lunghezza:

$$i_2 = \frac{a-b}{2}, \quad i_3 = \frac{a-b}{4}, \quad i_4 = \frac{a-b}{8}, \ldots$$

E quindi, in generale, possiamo scrivere:

$$i_n = \frac{a-b}{2^{n-1}}$$

Inoltre, possiamo dire che

$$f(a_n) > 0, \forall n \in \mathbb{N}$$

e

$$f(b_n) < 0, \forall n \in \mathbb{N}$$

per definizione (o costruzione, comunque la si voglia chiamare) delle due successioni.

Sicuramente poi

$$a_n < b_n, \forall n \in \mathbb{N}$$

Poichè le due successioni abbiamo detto che rappresentano gli estremi di un intervallo di lunghezza non nulla, e l'estremo inferiore in questo caso è sempre minore dell'estremo superiore.

Infine, a_n è una successione crescente (non strettamente), mentre b_n è una successione decrescente (non strettamente): per convincersene basta pensare a come costruiamo queste due successioni (e al fatto che il valore medio tra due numeri è sempre maggiore del più piccolo e minore del più grande). Quindi:

$$a_1 \leq a_2 \leq a_3 \leq \ldots$$

e

$$b_1 \geq b_2 \geq b_3 \geq \ldots$$

Combinando queste ultime due coppie di affermazioni, otteniamo che

$$a_n < b_n \leq b_1, \forall n \in \mathbb{N}$$

e

$$b_n > a_n \geq a_1, \forall n \in \mathbb{N}$$

Quindi abbiamo che la successione a_n è crescente e superiormente limitata, e quindi possiede limite in $\mathbb{R}$ (chiamiamolo α), e b_n è decrescente ed inferiormente limitata, e quindi lei pure possiede limite in $\mathbb{R}$ (chiamiamolo β). Sicuramente poi:

$$\alpha \leq \beta$$

Ma possiamo in realtà dire anche di più. Studiamo infatti la successione che determina la lunghezza degli intervalli:

$$b_n - a_n = i_n = \frac{a-b}{2^{n-1}} \to 0$$

Ma, se guardiamo le regole algebriche per ottenere il limite di una differenza di successioni con limite reale, ci accorgiamo che l'unico modo con cui può risultare 0 è che i due limiti siano uguali, quindi:

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n$$

ovvero

$$\alpha = \beta$$

Ormai la dimostrazione è giunta quasi al termine.

Per definizione di limite, noi sappiamo che

$$\lim_{n\to\infty} f(a_n)$$

è come dire

$$\lim_{x\to\alpha} f(x)$$

Ma, dato che $f(x)$ è continua, allora questo limite è uguale a

$$f(\alpha)$$

Quindi i vari valori che assume la f all'estremo sinistro tendono al valore $f(\alpha)$. Per b_n potremmo fare un ragionamento analogo, arrivando a dire che

$$\lim_{n\to\infty} f(b_n) = \beta$$

Ma $\beta = \alpha$, quindi:

$$\lim_{n\to\infty} f(a_n) = \lim_{n\to\infty} f(b_n) = f(\alpha)$$

Ora, dal primo limite che abbiamo usato per definire $f(\alpha)$, dato che tutti gli $f(a_n)$ sono positivi, come avevamo mostrato più sopra, otteniamo, per il teorema sulla permanenza del segno, che

$$f(\alpha) \geq 0$$

Mentre dal secondo limite, dato che tutti gli $f(b_n)$ sono negativi, abbiamo che

$$f(\beta) \leq 0 \quad \Longrightarrow \quad f(\alpha) \leq 0$$

Unendo queste due affermazioni, otteniamo che

$$f(\alpha) = 0$$

Ecco quindi che abbiamo trovato il nostro valore c anche nel caso in cui la progressione degli a_n e b_n non si fermi mai. CVD.

Altro importante teorema è il seguente:

Teorema 5.3.2 (Teorema del valore intermedio (o di Bolzano)) *Se $f : I \subseteq \mathbb{R} \to \mathbb{R}$ è continua (I indica un intervallo), o f è costante, o $f(I)$ è un intervallo a sua volta.*

Importante inanzitutto definire bene cosa intende per intervallo: un intervallo è un insieme numerico che possiede questa proprietà: se $a, b \in I$, allora $[a, b] \in I$.

Per dimostrare questo teorema, prendiamo inanzitutto due qualsiasi punti di $f(I)$: y_1 ed y_2. Per dimostrare il teorema, quel che dobbiamo fare è quindi dimostrare che, qualunque coppia di punti y_1 ed y_2 appartenendi ad $f(I)$ essi siano, $[y_1, y_2] \in f(I)$.

Prendiamo quindi un qualunque $y \in [y_1, y_2]$. Dato che $y_1 \in f(I)$, allora sarà immagine di un certo x_1, e così anche y_2 sarà immagine di un x_2. Sicuramente $x_1 \neq x_2$, dato che $y_1 \neq y_2$ (se fosse altrimenti, avremmo uno stesso punto con due immagini differenti, cosa che non può essere dato che f è una funzione). Quindi possiamo prendere l'intervallo $[x_1, x_2]$, il quale apparterrà sicuramente ad I, dato che I è un intervallo.

Definiamo ora così una nuova funzione, $g(x)$, che ci sarà utile:

$$g(x) \stackrel{def}{=} f(x) - y$$

Abbiamo che

$$g(x_1) = f(x_1) - y = y_1 - y < 0$$

(infatti $y \in [y_1, y_2]$), e

$$g(x_2) = f(x_2) - y = y_2 - y > 0$$

(per la stessa ragione). Quindi, per il teorema degli zeri, $\exists c \in [x_1, x_2] : g(c) = 0$. Ma se $g(c) = 0$, allora è come dire che $f(c) - y = 0$ (per definizione di $g(x)$), ovvero $y = f(c)$. Quindi y è effettivamente immagine di un qualche punto di I, e dato che l'unica supposizione che abbiamo fatto su y è che appartenga all'intervallo $[y_1, y_2]$, allora questa affermazione varrà per qualunque y appartenente a quell'intervallo. CVD.

Ultimo, fondamentale, teorema, che qui non dimostreremo, è il seguente:

Teorema 5.3.3 (Teorema di Weierstrass) *Se $f : [a, b] \to \mathbb{R}$, allora f ha massimo e minimo assoluti (ovvero l'insieme $\{f(x) \mid x \in [a, b]\}$ possiede massimo e minimo).*

Esempio 5.3.1 *Data l'equazione*

$$x^3 - 3x + 3 = 0$$

Dimostrare che possiede almeno una soluzione reale

Chiaramente non possiamo utilizzare il teorema fondamentale dell'algebra, perchè in quel caso potrei parlare solo di soluzioni complesse. Però possiamo osservare che, ad esempio, nel punto 0 il polinomio

$$f(x) = x^3 - 3x + 3$$

vale 3, quidi un valore positivo. Nel punto -3 invece vale -15, ovvero un valore negativo: per il teorema degli zeri, quindi, in almeno un punto di questo intervallo il polinomio si annulla, e quindi ha una soluzione l'equazione di partenza.

Esempio 5.3.2 *Dimostrare che l'equazione*

$$x^5 - \frac{\sqrt{2}}{12}x^4 + 315x^3 + e^{13/2}x^2 - x + 4\pi = 0$$

ha almeno una soluzione.

Nonostante l'apparente enorme complessità della funzione di partenza (chiamiamola come al solito $f(x)$), si può dimostrare senza andare a fare calcoli complessi la veridicità dell'affermazione proposta.

Infatti, esaminando semplicemente il grado (dispari) e il coefficiente del monomio di grado maggiore della funzione (ovvero di x^5, che ha per coefficiente 1, un valore positivo), possiamo già dire che:

$$\lim_{x \to -\infty} f(x) = -\infty$$

e

$$\lim_{x \to +\infty} f(x) = +\infty$$

Ora, dato che il dominio ($\mathbb{R}$) è un intervallo, anche $f(I)$ deve essere un intervallo. E dato che la funzione deve poter assumere valori grandi in positivo e negativo quanto uno vuole (a causa dei due limiti espressi precedentemente), allora deve poter assumere, in particolare, almeno un valore positivo ed uno negativo – tra questi due valori è quindi contenuta sicuramente la soluzione per il teorema degli zeri, dato che la funzione è continua.

Questo ragionamente, ovviamente, si può fare genericamente per qualunque polinomio di grado dispari.

5.4 Limiti notevoli

Esistono alcuni limiti, espressi sotto forma indeterminata, che si rivelano di particolare importanza nel calcolo degli altri limiti e, vedremo in seguito, delle derivate – tali limiti vengono detti "limiti notevoli".

$$\lim_{x \to 0} \frac{\sin x}{x} = 1 \quad \left[\frac{0}{0}\right]$$

Prendiamo, come in figura, al momento solo il primo quadrante della circoferenza goniometrica (e quindi gli x compresi tra 0 e $\pi/4$ - per i valori negativi, il discorso non cambia).

Prendiamo l'area del triangolo OAB, e calcoliamone l'area:

$$A_{OAB} = \frac{bh}{2} = \frac{\overline{OB} \cdot \sin x}{2} = \frac{\sin x}{2}$$

L'area invece dell'arco OAB è:

$$A = \frac{x \cdot r^2}{2} = \frac{x}{2}$$

Infine, l'area del triangolo OCB è

$$A_{OCB} = \frac{\overline{OB} \cdot \overline{BC}}{2} = \frac{\tan x}{2}$$

Ora, dalla figura è evidente che

$$A_{OAB} < A < A_{OCB}$$

quindi

$$\frac{\sin x}{2} < \frac{x}{2} < \frac{\tan x}{2}$$

Da cui ricavo, moltiplicando tutti i membri per $2/\sin x$:

$$1 < \frac{x}{\sin x} < \frac{1}{\cos x}$$

E, invertendo:

$$1 > \frac{\sin x}{x} > \cos x$$

Per la regola del confronto (la stessa delle successioni, che è facile intuire che vale anche per i limiti di funzione reale), otteniamo che 1, essendo una costante, per $x \to 0$, tende ad 1, $\cos x$, essendo una funzione continua, tende a $\cos 0 = 1$, e quindi $\sin x/x$ è "costretto" a tendere ad 1 pure lui. CVD.

Un altro modo per scrivere un limite del genere, è il seguente:

$$\sin x \sim x, \text{ per } x \to 0$$

Che ha lo stesso significato che per le successioni. Inoltre, in questo caso abbiamo anche un significato più immediato, ovvero questo confronto asintotico ci dice che $\sin x$ ed x si comportano in modo simile in un intorno piccolo del punto 0, e che il loro comportamento è sempre più simile man mano che si avvicinano a 0.

$$\lim_{x \to 0} \frac{1 - \cos x}{x} = 0$$

Con semplici trasformazioni:

$$\frac{1 - \cos x}{x} = \frac{(1 - \cos x)(1 + \cos x)}{x(1 + \cos x)} = \frac{\sin^2 x}{x(1 + \cos x)} = \left(\frac{\sin x}{x}\right)^2 \frac{x}{1 + \cos x}$$

Ora, per $x \to 0$, il rapporto $\sin x/x$, come abbiamo appena dimostrato, tende ad 1, ed il suo quadrato quindi pure, mentre $x \to 0$ e $1 + \cos x \to 2$, perchè funzioni continue. Il limite nel suo complesso tende quindi a $1 \cdot 0/2 = 0$.

In questo caso il limite trovato non ci aiuta a stabilire un confronto asintotico – uno simile, tuttavia, lo fa, ed è il seguente:

$$\lim_{x \to 0} \frac{1 - \cos x}{x^2} = \frac{1}{2}$$

In questo caso, attraverso passaggi simili ai precedenti, arriviamo alla forma:

$$\left(\frac{\sin x}{x}\right)^2 \cdot \frac{1}{1 + \cos x}$$

che, non avendo più la x al numeratore, da come risultato $1/2$, per l'appunto.

Con qualche calcolo, si arriva facilmente ad ottenere una nuova forma asintotica:

$$\cos x \sim 1 - \frac{x^2}{2}$$

Sempre per $x \to 0$.

L'unico limite che ci limiteremo ad indicare senza dimostrare è il seguente:

$$\lim_{x \to \pm\infty} \left(1 + \frac{1}{x}\right)^x = e$$

La cui dimostrazione si basa, ovviamente, sul fatto che la serie corrispondente, ma che non è altrettanto facile, dato che, per *qualunque* successione $x_n \to \pm\infty$ deve valere:

$$\left(1+\frac{1}{x_n}\right)^{x_n} \to e$$

fatto dimostrabile per $x_n = n$, ma ancora da provare per altri generi di successioni.

Se ora calcoliamo il logaritmo di questo ultimo limite notevole, otteniamo per il membro di sinistra il valore:

$$\log\left(1+\frac{1}{x}\right)^{x} = x\log\left(1+\frac{1}{x}\right)$$

E per il membro di destra:

$$\log e = 1$$

Quindi possiamo dire che

$$\lim_{x\to\pm\infty} x\log\left(1+\frac{1}{x}\right) = 1$$

Poiché il logaritmo è una funzione continua. Ora, ponendo $t = 1/x$, possiamo anche dire:

$$x\log\left(1+\frac{1}{x}\right) = \frac{\log\left(1+\frac{1}{x}\right)}{\frac{1}{x}} = \frac{\log(1+t)}{t}$$

Ora, per $x \to \pm\infty$, $t \to 0$, quindi, possiamo anche scrivere, tornando alla normale variabile x:

$$\lim_{x\to 0}\frac{\log(1+x)}{x} = 1$$

Ecco quindi che abbiamo trovato un altro utile confronto asintotico.

Infine, ponendo $x = e^t - 1$, abbiamo:

$$\frac{\log(1+x)}{x} = \frac{\log(1+e^t-1)}{e^t-1} = \frac{t}{e^t-1}$$

Inoltre abbiamo che, per $x \to 0$, ovviamente $e^t - 1 \to 0$, ovvero $t \to 0$. Quindi:

$$\lim_{x\to 0}\frac{e^x-1}{x} = 1$$

Ed anche questo può essere espresso come confronto asintotico.

Riassumendo tutti i limiti notevoli per ora trovati, abbiamo:

- Per $x \to 0$, $\quad \sin x \sim x$
- $\lim_{x\to 0}\dfrac{1-\cos x}{x} = 0$
- Per $x \to 0$, $\quad \cos x \sim 1 - \dfrac{x^2}{2}$
- $\lim_{x\to\pm\infty}\left(1+\dfrac{1}{x}\right)^{x} = e$
- Per $x \to 0$, $\quad \log(1+x) \sim x$
- Per $x \to 0$, $\quad e^x - 1 \sim x$

Come si nota siamo quasi sempre riusciti a stabilire delle approssimazioni asintotiche per $x \to 0$, ed in questo caso sempre con dei polinomi in x, molto più facili da trattare delle corrispondenti funzioni.

Capitolo 6

Derivata

6.1 Definizione di Derivata

Data una funzione

$$f : I \to \mathbb{R}$$

continua sull'intervallo I, so che

$$\lim_{x \to x_0} f(x) = f(x_0)$$

Ma una domanda che può sorgere spontanea è: a che "velocità" tende a tale valore questa funzione?

Cominciamo col vedere che un valore associato a questo concetto potremo ricavarlo prendendo vari punti x sempre più vicini ad x_0 (quindi, $x \to x_0$), ed esaminando le due grandezze $f(x) - f(x_0)$ ed $x - x_0$. La prima rappresenta la distanza verticale tra i due punti individuati, e la seconda la distanza orizzontale. Una stima del valore che cerchiamo sarà il loro rapporto: è tanto maggiore quanto maggiore è la distanza che la funzione compie per arrivare ad x_0. Il rapporto:

$$\frac{f(x) - f(x_0)}{x - x_0}$$

viene detto *rapporto incrementale*. Il suo limite per $x \to x_0$ (chiaramente una forma indeterminata del tipo $0/0$), viene chiamata *derivata della funzione nel punto* x_0 e si indica nei seguenti modi:

$$Df(x), \quad D_x f, \quad f'(x), \quad \frac{\mathrm{d}f}{\mathrm{d}x}$$

Quindi:

$$Df(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

Il generale, la funzione derivata è quindi così definita.

Una funzione per la quale, nel punto x_0, tale limite è definito, viene detta *derivabile* nel punto x_0, ed una funzione derivabile in ogni punto del suo dominio viene detta, semplicemente, *derivabile*.

Si scopre che *tutte le funzioni elementari sono derivabili sul loro dominio*. Passiamo ora ad esaminare tali derivate per queste funzioni:

6.2 Derivate delle funzioni elementari

6.2.1 Funzione costante

$$f(x) = c, c \in \mathbb{R} \qquad I = \mathbb{R}, \text{ed ivi continua}$$

Calcoliamo il rapporto incrementale:

$$\frac{f(x) - f(x_0)}{x - x_0} = \frac{c - c}{x - x_0} = 0, \quad x \neq x_0$$

Ovviamente, il limite di tale funzione è 0 – la condizione $x \neq x_0$ è rispettata poichè nella definizione di limite la successione degli a_n non deve mai essere uguale a x_0. Quindi abbiamo che:

$$f'(x_0) = 0$$

6.2.2 Funzione identità

$$f(x) = x, \qquad I = \mathbb{R}, \text{ed ivi continua}$$

Il rapporto incrementale vale:

$$\frac{f(x) - f(x_0)}{x - x_0} = \frac{x - x_0}{x - x_0} = 1$$

Il cui limite è ovviamente 1, quindi:

$$f'(x_0) = 1$$

6.2.3 Funzione x^2

$$f(x) = x^2, \qquad I = \mathbb{R}, \text{ed ivi continua}$$

Il suo rapporto incrementale è:

$$\frac{f(x) - f(x_0)}{x - x_0} = \frac{x^2 - x_0^2}{x - x_0} = \frac{(x - x_0)(x + x_0)}{x - x_0} = x + x_0$$

Il cui limite, per $x \to x_0$ è:

$$x_0 + x_0 = 2x_0$$

Quindi:

$$f'(x_0) = 2x_0$$

6.2.4 Funzione x^n

$$f(x) = x^n, n \in \mathbb{N} \qquad I = \mathbb{R}, \text{ed ivi continua}$$

Il suo rapporto incrementale è:

$$\frac{f(x) - f(x_0)}{x - x_0} = \frac{x^n - x_0^n}{x - x_0} =$$

$$\frac{(x - x_0)(x^{n-1} + x^{n-2}x_0 + \cdots + xx_0^{n-2} + x_0^{n-1})}{x - x_0} =$$

$$x^{n-1} + x^{n-2}x_0 + \cdots + xx_0^{n-2} + x_0^{n-1}$$

Ora, questi sono esattamente n termini. Per $x \to x_0$, abbiamo:

$$\underbrace{x_0^{n-1} + x_0^{n-2}x_0 + \cdots + x_0x_0^{n-2} + x_0^{n-1}}_{\text{n volte}} = nx_0^{n-1}$$

Quindi:

$$f'(x_0) = nx_0^{n-1}, n \in \mathbb{N}$$

Si può notare come questa uguaglianza valga, e quindi riassuma, anche quelle precedentemente dimostrate.

6.2.5 Funzione e^x

$$f(x) = e^x$$

Il rapporto incrementale:

$$\frac{f(x) - f(x_0)}{x - x_0} = \frac{e^x - e^{x_0}}{x - x_0} = \frac{e^{x_0 + (x - x_0)} - e^{x_0}}{x - x_0} = \frac{e^{x_0}e^{x - x_0} - e^{x_0}}{x - x_0}$$

Raccogliendo e^{x_0}:

$$e^{x_0} \cdot \frac{e^{x - x_0} - 1}{x - x_0}$$

Ponendo $t = x - x_0$, abbiamo che per $x \to x_0$, $t \to 0$, e quindi abbiamo per il secondo fattore un limite notevole che tende ad 1: il limite dell'intero rapporto incrementale è quindi

$$f'(x_0) = e^{x_0}$$

La funzione esponenziale a base e, come vedremo, è l'unica ad avere per derivata sè stessa – una proprietà che si rivelerà importantissima:

$$(e^x)' = e^x$$

6.2.6 Funzione seno

$$f(x) = \sin x$$

Il rapporto incrementale:

$$\frac{f(x) - f(x_0)}{x - x_0} = \frac{\sin(x_0 + (x - x_0)) - \sin x_0}{x - x_0}$$

Ora, ricordando che:

$$\sin(\alpha + \beta) = \sin\alpha\cos\beta + \cos\alpha\sin\beta$$

Possiamo scrivere $\sin(x_0 + (x - x_0))$ per esteso:

$$\frac{\sin x_0 \cos(x - x_0) + \cos x_0 \sin(x - x_0) - \sin x_0}{x - x_0} =$$

$$\sin x_0 \frac{\cos(x - x_0) - 1}{x - x_0} + \cos x_0 \frac{\sin(x - x_0)}{x - x_0}$$

Ma, se prendiamo $t = x - x_0$, per $x \to x_0$, $t \to 0$, e quindi la frazione al primo addendo è un limite notevole che tende a 0, mentre quella al secondo addendo è sempre un limite notevole, ma che questa volta tende ad 1. Il risultato finale è quindi assai semplice:

$$f'(x_0) = \cos x_0$$

6.2.7 Funzione coseno

$$f(x) = \cos x$$

Il rapporto incrementale vale:

$$\frac{f(x) - f(x_0)}{x - x_0} = \frac{\cos x - \cos x_0}{x - x_0}$$

Con passaggi analoghi a quelli precedenti, e ricordando che:

$$\cos(\alpha + \beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta$$

Otteniamo:

$$\cos x_0 \frac{\cos(x - x_0) - 1}{x - x_0} - \sin x_0 \frac{\sin(x - x_0)}{x - x_0}$$

Con cambiamenti di variabili identici a quelli precedenti, abbiamo che il primo addendo tende a $\cos x_0 \cdot 0 = 0$ ed il secondo a $-\sin x_0 \cdot 1 = -\sin x_0$, che è per l'appunto la derivata cercata, anche in questo caso piuttosto semplice:

$$f'(x_0) = -\sin x_0$$

6.2.8 Funzione logaritmo

$$f(x) = \log x$$

Rapporto incrementale:

$$\frac{f(x) - f(x_0)}{x - x_0} = \frac{\log x - \log x_0}{x - x_0} =$$

$$= \frac{\log(x/x_0)}{x - x_0} = \frac{\log(x/x_0)}{(x/x_0 - 1)\, x_0} =$$

$$= \frac{1}{x_0} \cdot \frac{\log\left(1 + (x/x_0 - 1)\right)}{x/x_0 - 1}$$

A questo punto, se sostituiamo ad $x/x_0 - 1$ la variabile t (che tende a 0 per x che tende ad x_0), abbiamo, al limite:

$$\lim_{x \to x_0\ (t \to 0)} \frac{1}{x_0} \frac{\log(1 + t)}{t}$$

il cui secondo fattore è uno dei limiti notevoli già studiato, e che tende ad 1; otteniamo quindi per derivata:

$$f'(x_0) = \frac{1}{x_0}$$

6.3 Continuità e derivabilità

Che rapporto c'è tra continuità e derivabilità? Abbiamo visto ad esempio che tutte le funzioni elementari sono sia continue che derivabili sul loro dominio, e quindi verrebbe da pensare che i due concetti siano equivalenti, ma ciò non è così : una funzione derivabile è anche continua, ma non il contrario, come possono mostrare i seguenti esempi:

Esempio 6.3.1 *Dimostrare che la funzione*

$$f(x) = |x|$$

non è derivabile nel punto 0 *pur essendo continua*

Che la funzione sia continua nel punto 0 *è facile da mostrare.* $f(0) = 0$*, ed il limite:*

$$\lim_{x\to 0} |x|$$

È facile mostrare che vale 0, esaminando limite destro e sinistro separatamente. Per quando riguarda invece la derivabilità nel punto 0, bisogna esaminare l'esistenza o meno del limite:

$$\lim_{x\to 0} \frac{|x| - |0|}{x - 0} = \lim_{x\to 0} \frac{|x|}{x}$$

Ma la funzione di cui calcoliamo il limite, per x *negativi, vale* -1*, mentre per* x *positivi vale* 1 *– avendo limite destro e sinistro diversi, si può concludere che tale limite non esiste, e che la* $f(x)$ *non è quindi derivabile nel punto 0.*

Esempio 6.3.2 *Dimostrare che la funzione*

$$f(x) = \sqrt{|x|}$$

Non è derivabile nel punto 0

Anche in questo caso ci troviamo di fronte ad una funzione continua ($|x|$ *abbiamo già mostrato prima essere continua, e la radice applicata su valori positivi o nulli è continua anch'essa). Calcoliamo ora il rapporto incrementale:*

$$\frac{\sqrt{|x|} - \sqrt{0}}{x - 0} = \frac{\sqrt{|x|}}{x} = \begin{cases} \dfrac{1}{\sqrt{x}}, & \text{per } x \geq 0 \\ -\dfrac{1}{\sqrt{-x}}, & \text{per } x \leq 0 \end{cases}$$

Quindi, il limite sinistro della funzione per $x \to 0$ *vale* $-\infty$*, mentre quello destro vale* $+\infty$*. Quindi, il rapporto incrementale non possiede limite, e la funzione in* 0 *non è derivabile.*

6.4 Interpretazione geometrica della derivata

Per poter capire cosa rappresenta geometricamente la derivata, procediamo per gradi. Cominciamo intanto col capire il significato del rapporto incrementale: se noi prendiamo il punto x_0 ed un altro $x \neq x_0$, definiamo i due punti $(x_0, f(x_0))$ e $(x, f(x))$. La retta passante per questi due punti sarà secante alla funzione in almeno due punti (quelli corrispondenti all'ascissa x ed x_0 per l'appunto). Il suo coefficiente angolare sarà:

$$m = \frac{\Delta y}{\Delta x} = \frac{f(x) - f(x_0)}{x - x_0}$$

Che è per l'appunto il rapporto incrementale! Ora, la derivata è il limite del rapporto incrementale per $x \to x_0$ - geometricamente, questo significa che la derivata è il limite a cui tendono i coefficienti angolari delle varie rette che passano per $(x_0, f(x_0))$ ed $(x, f(x))$, man mano che x si avvicina ad x_0. È facile intuire che si avranno sempre delle secanti in punti, ovviamente, sempre più vicini, fintanto che x ed x_0 non saranno lo stesso punto. Quindi avremo una retta che è secante in un unico punto (x_0) al grafico della funzione - ma questo significa che tale retta è *tangente* in tale punto!

Ecco quindi trovato il significato geometrico di derivata: la derivata in x_0 di $f(x)$ è il coefficiente angolare della retta tangente al grafico di $f(x)$ nel punto x_0.

L'equazione della retta tangente si ricava facilmente, ed è:

$$y = f(x_0) + f'(x_0)(x - x_0)$$

Quindi, avere una funzione *continua* in x_0 significa che essa in x_0 non possiede "strappi" o "balzi", invece dire che è *derivabile* significa che nel punto x_0 possiede tangente. Ovviamente, per possedere tangente in un punto, il grafico di una funzione dovrà anche essere continuo, però il semplice fatto che sia continuo in x_0 non permette di dire che lì sarà anche possibile costruire la tangente – ad esempio nei due esempi precedenti avevamo nel punto 0 una "punta" (o "*cuspide*"), dove la funzione era sicuramente continua, ma non era possibile costruire nessuna retta tangente (e infatti lì la funzione non era derivabile).

6.5 Regole di calcolo delle derivate

Abbiamo già le derivate di tutte le funzioni elementari – ora, se riusciamo anche a trovare le derivate delle possibile "combinazioni" di funzioni a partire dalle derivate dei suoi componenti elementari potremo quindi calcolare la derivata di ogni funzione. I possibili operatori che si possono applicare su una funzione sono quelli aritmetici (+,-,*, /), in aggiunta la composizione di funzioni ($f \circ g$) e le funzioni inverse (f^{-1}).

6.5.1 Somma e differenza di funzioni

Qui la regola è semplicissima:

$$D(f \pm g) = Df \pm Dg$$

Ovvero: la derivata di una somma è la somma delle derivate

6.5.2 Prodotto di una funzione per una costante

Anche qui la regola è assai intuitiva:

$$D(kf) = kDf, \quad k \in \mathbb{R}$$

Ovvero: la derivata del prodotto di una costante per una funzione è uguale al prodotto della costante per la derivata della funzione – o, detto in modo più semplice, si può far uscire una costante che moltiplica una funzione dal segno di derivata.

A questo punto siamo già in grado, ad esempio, di calcolare la derivata di un qualunque polinomio.

Esempio 6.5.1 *Calcolare*

$$D(3x^5 + 5x^3 - 4x^2 + 3)$$

Inanzitutto, questa è la derivata di una somma di funzioni, ed è quindi uguale alla somma delle derivate:

$$D(3x^5) + D(5x^3) - D(4x^2) + D(3) =$$

Qui abbiamo invece vari prodotti di una costante per una funzione, quindi facciamo uscire le costanti – inoltre, sappiamo già calcolare la derivata della funzione (costante) 3, che è 0, e che quindi rimuoviamo dalla somma:

$$= 3D(x^5) + 5D(x^3) - 4D(x^2) =$$

Qui abbiamo tutte derivate di potenze, che sappiamo calcolare

$$= 15x^4 + 15x^2 - 8x$$

che è, come si può notare, un polinomio di quarto grado. In generale: la derivata di un polinomio di grando n è un polinomio di grado $n - 1$.

Esempio 6.5.2 *Calcolare*

$$D(e^x + 2\log x)$$

I successivi passaggi da applicare sono:

$$D(e^x) + D(2\log x) = e^x + 2D(\log x) = e^x + 2 \cdot \frac{1}{x} = e^x + \frac{2}{x}$$

6.5.3 Prodotto di funzioni

Qui le regole cominciano a complicarsi: infatti la derivata di un prodotto *non* è il prodotto delle derivate! Invece, vale la seguente regola:

$$D(f \cdot g) = f \cdot D(g) + D(f) \cdot g$$

Che, generalizzata ad n funzioni, si può facilmente dimostrare, per mezzo della regola appena mostrata, che diventa:

$$D(f_1 \cdot f_2 \cdot f_3 \cdot \dots) = D(f_1) \cdot f_2 \cdot f_3 \cdot \dots + f_1 \cdot D(f_2) \cdot f_3 \cdot \dots + f_1 \cdot f_2 \cdot D(f_3) \cdot \dots + \dots$$

Per dimostrare la prima affermazione, si deve partire dalla definizione di derivata – calcoliamo quindi il rapporto incrementale di $f(x)g(x)$:

$$\frac{f(x)g(x) - f(x_0)g(x_0)}{x - x_0} = \frac{f(x)g(x) - f(x_0)g(x) + f(x_0)g(x) - f(x_0)g(x_0)}{x - x_0}$$

Abbiamo aggiunto e tolto la stessa quantità, quindi non abbiamo modificato il valore di questa funzione. Con alcuni raccoglimenti, otteniamo:

$$\frac{(f(x)-f(x_0))g(x)+(g(x)-g(x_0))f(x_0)}{x-x_0}$$

Che, spezzando in due parti la somma e ponendo davanti i fattori comuni, diventa:

$$\frac{f(x)-f(x_0)}{x-x_0}g(x)+\frac{g(x)-g(x_0)}{x-x_0}f(x_0)$$

Ecco ora che abbiamo finalmente messo in evidenza i rapporti incrementali delle due funzioni separatamente. Se quindi ora passiamo al limite $x \to x_0$ al primo e ultimo membro (che, essendo uguali, hanno uguali anche i rispettivi limiti), otteniamo:

$$D(f \cdot g) = D(f)g + fD(g)$$

C'è da notare che $\lim_{x\to x_0} g(x) = g(x_0)$ è stato possibile solo perchè $g(x)$ è derivabile, e quindi anche continua.

Esempio 6.5.3 *Calcolare*

$$D\left[\left(2x^2+3\right)e^x\right]$$

Essendo il rapporto di due derivate, otteniamo che il risultato sarà:

$$D(2x^2+3)\cdot e^x + D(e^x)\cdot(2x^2+3) = 4xe^x + e^x(2x^2+3) = (2x^2+4x+3)e^x$$

6.5.4 Reciproco di una funzione

Anche in questo la regola è più complessa di quanto ci si potrebbe aspettare. Presa un funzione

$$g(x) \neq 0$$

la derivata del suo rapporto vale:

$$D\frac{1}{g} = -\frac{Dg}{g^2}$$

E si dimostra in modo analogo a prima, partendo dal rapporto incrementale:

$$\frac{1/g(x)-1/g(x_0)}{x-x_0} = \frac{1}{g(x)g(x_0)}\cdot\frac{g(x_0)-g(x)}{x-x_0}$$

Ora abbiamo quasi evidenziato il rapporto incrementale di g, che però ha segno sbagliato – per risolvere il problema, mettiamo un meno davanti (e già che ci siamo passiamo subito al limite):

$$\lim_{x\to x_0}\frac{1}{g(x)g(x_0)}\cdot\left[-\frac{g(x)-g(x_0)}{x-x_0}\right] = \frac{1}{g(x)^2}\cdot[-Dg(x)] = -\frac{Dg}{g^2}$$

Ecco dimostrata la regola.

6.5.5 Rapporto di funzioni

È semplicemente la combinazione delle due regole appena viste:

$$D\frac{f}{g} = D\left(f\cdot\frac{1}{g}\right) = Df\frac{1}{g} + fD\frac{1}{g} = \frac{Df}{g} - \frac{fDg}{g^2} = \frac{gDf-fDg}{g^2}$$

Esempio 6.5.4 *Calcolare:*

$$D\csc x$$

Ricordando che $\csc x$ *(cosecante di* x*) è*

$$\frac{1}{\sin x}$$

Quindi si tratta di trovare la derivata di un reciproco:

$$D\frac{1}{\sin x} = -\frac{D\sin x}{\sin^2 x} = -\frac{\cos x}{\sin^2 x} = -\csc x\cot x$$

($\cot x$*, o cotangente di* x*, è il reciproco della tangente).*

Esempio 6.5.5 *Calcolare:*

$$D\sec x$$

Ricordando che $\sec x$ *(secante di* x*) è*

$$\frac{1}{\cos x}$$

Anche qui si tratta di trovare la derivata di un reciproco:

$$D\frac{1}{\cos x} = -\frac{-\sin x}{\cos^2 x} = \sec x \tan x$$

6.5.6 Funzioni composte (regola di catena)

La derivata di una funzione composta è:

$$D(f(g(x))) = Df(g) \cdot Dg(x)$$

ovvero è la derivata della prima funzione utilizzando la seconda come variabile indipendente moltiplicata per la derivata della seconda. Come suggerisce il nome della regola, si può estendere questa regola ad un qualunque numero di funzioni:

$$D(f(g(h(j(\ldots))))) = D(f(g)) \cdot D(g(h)) \cdot D(h(j)) \cdot \ldots$$

L'espressione di ciò scritto con la notazione differenziale è forse anche più facile da ricordare:

$$\frac{\mathrm{d}f(g(x))}{\mathrm{d}x} = \frac{\mathrm{d}f}{\mathrm{d}g} \cdot \frac{\mathrm{d}g}{\mathrm{d}x}$$

Come se il dg si semplificasse tra le due frazioni – anche per più funzioni il discorso non cambia:

$$D(f(g(h(j(\ldots))))) = \frac{\mathrm{d}f}{\mathrm{d}g} \cdot \frac{\mathrm{d}g}{\mathrm{d}h} \cdot \frac{\mathrm{d}h}{\mathrm{d}j} \cdot \ldots$$

Esempio 6.5.6 *Calcolare*

$$D\left(e^{\frac{x^2-1}{x^2+1}}\right)$$

Abbiamo una funzione composta, quella più interna è un rapporto di polinomi, mentre quella più esterna è un'esponenziale; quindi, risolvendo con la regola della catena, ottendiamo:

$$\frac{d\, e^{\frac{x^2-1}{x^2+1}}}{d\, \frac{x^2-1}{x^2+1}} \cdot D\frac{x^2-1}{x^2+1}$$

Nel primo fattore il rapporto tra differenziali dice, in pratica, che dobbiamo derivare l'esponenziale utilizzando come variabile indipendente l'intero blocco del rapporto di polinomi, mentre il secondo fattore è la semplice derivata di un rapporto, quindi otteniamo:

$$e^{\frac{x^2-1}{x^2+1}} \cdot \frac{2x(x^2+1) - 2x(x^2-1)}{(x^2+1)^2}$$

Che, con le dovute semplificazioni, arriva ad essere:

$$e^{\frac{x^2-1}{x^2+1}} \frac{4x}{(x^2+1)^2}$$

Esempio 6.5.7 *Calcolare*

$$Da^x, \quad a \neq 1$$

Ricordiamo che

$$\ln a^x = x \ln a$$

E quindi

$$a^x = e^{\ln a^x} = e^{x \ln a}$$

La cui derivata, essendo una funzione composta, risulta essere:

$$\frac{d\, e^{x \ln a}}{x \ln a} \cdot D(x \ln a) = e^{x \ln a} \ln a = a^x \ln a$$

Ecco quindi trovata la derivata di una esponenziale a base qualunque.

Esempio 6.5.8 *Calcolare*

$$D\log_a x, \quad x > 0, x \neq 1$$

Ricordando che

$$\log_a b \cdot \log_b c = \log_a c$$

abbiamo che

$$\log_a x = \log_a e \cdot \ln x$$

La cui derivata vale

$$D_x(\log_a e \cdot \ln x) = \log_a e \, D(\ln x) = \frac{1}{x \ln a}$$

Ed ecco quindi la regola di derivazione di un logaritmo di base qualunque.

6.5.7 Funzione inverse

L'ultimo caso che ci rimane da esaminare è quello delle funzioni inverse; il problema in pratica è: sapendo la derivata di una funzione, come si trova quella della funzione inversa? Già nei paragrafi sulle funzioni continue avevamo dato le condizioni necessarie per l'invertibilità di una funzione. Supponiamo ora che f^{-1} sia derivabile, e prendiamo per dominio di f l'intervallo I e per sua immagine l'intervallo J (è un intervallo per il teorema di Bolzano). Ora prendiamo un punto $x \in J$, ovvero nel dominio di f^{-1}. Quindi, $f^{-1}(x) \in I$. In questo caso, possiamo applicare a questo risultato la f, e otteniamo di nuovo x: abbiamo fatto in pratica un salto avanti ed uno indietro percorrendo la stessa strada prima in un verso e poi in quello opposto: ovvero prima nel verso della funzione inversa e poi della funzione di base. Ciò si esprime con questa uguaglianza:

$$f(f^{-1}(x)) = x, \quad \forall x \in J$$

Essendo vera questa affermazione per ogni x di J, applichiamo la derivata ad entrambi i membri ed avremo ancora un'affermazione vera $\forall x \in J$:

$$D[f(f^{-1}(x))] = Dx$$

Ma

$$Dx = 1$$

Mentre l'altro membro è la derivata di una funzione composta, e quindi otteniamo:

$$f'(f^{-1}(x)) \cdot (f^{-1})'(x) = 1$$

Dividiamo ora entrambi i membri per $f'(f^{-1}(x))$, e quindi lo presupponiamo non nullo, ed otteniamo:

$$(Df^{-1})(x) = \frac{1}{(Df)(f^{-1}(x))}$$

Ovvero: la derivata dell'inversa di una funzione è l'inverso della derivata della funzione applicata alla funzione inversa della variabile indipendente.

Se facciamo qualche considerazione, vediamo che applichiamo sempre funzioni a variabili del dominio giusto: x appartiene a J, come detto all'inizio, e noi effettivamente al membro di sinistra applichiamo la derivata della funzione inversa, che ha dominio in J, alla x. Al membro destro, applichiamo prima la funzione inversa alla x, ottenendo quindi un risultato che sta in I, e quindi vi applichiamo la derivata della funzione, che effettivamente ha dominio in I.

Esempio 6.5.9 *Calcolare*

$$D(\arctan x)$$

Cominciamo col definire la funzione arcotangente *(*$\arctan x$*). Se prendiamo la funzione tangente, vediamo che essa, nel suo complesso, non può essere considerata invertibile, poiché non è biunivoca in tutto il suo intervallo di definizione (che è poi* $\mathbb{R} - \{\pi/2 + k\pi \mid k \in \mathbb{Z}\}$*). Tuttavia, se prendiamo un intervallo più piccolo, possiamo definire tale funzione: questo intervallo è uno degli infiniti intervalli* $(-\pi/2 + k\pi; \pi/2 + k\pi), k \in \mathbb{Z}$*. Infatti in questo intervallo la tangente è strettamente crescente e continua, e quindi senz'altro biunivoca. Per convenzione si prende l'intervallo* $(\pi/2; \pi/2)$*. L'immagine di tale dominio è* $\tan((-\pi/2; \pi/2)) = \mathbb{R}$*. Quindi, la funzione arcotangente avrà per dominio* $\mathbb{R}$ *e per codominio* $(-\pi/2; \pi/2)$*. Non è quindi giusto dire che l'arcotangente è la funzione inversa della tangente – è giusto invece dire che è la funzione inversa della tangente sull'intervall* $(-\pi/2; \pi/2)$*. Calcoliamo ora la derivata dell'arcotangente su tale intervallo:*

$$D(\arctan x) = \frac{1}{(D\tan)(\arctan x)} = \frac{1}{1 + \tan^2(\arctan x)} = \frac{1}{1 + x^2}$$

Dobbiamo però considerare il caso in cui $(D\tan)(\arctan x) = 0$*. Siamo fortunati, poichè è facile verificare che ciò non avviene mai.*

In questo caso il calcolo è stato molto facile - è bastato scegliere adeguatamente tra le due forme con cui si può esprimere la derivata dell'arcotangente. è inoltre interessante il fatto che l'arcotangente, come il logaritmo, pur essendo funzioni irrazionali*, hanno per derivate funzioni* razionali*.*

Esempio 6.5.10 *Calcolare*

$$D(\arcsin x)$$

Anche nel caso dell'arcoseno si fanno ragionamenti analoghi a quelli fatti per l'arcotangente. Anche in questo caso l'intervallo che si prende è $(-\pi/2;\pi/2)$, ma stavolta il dominio dell'arcoseno risulta essere $[-1;1]$. Calcoliamo ora la derivata richiesta:

$$D(\arcsin x) = \frac{1}{(D\sin)(\arcsin x)} = \frac{1}{\cos\arcsin x}$$

A questo punto bisogna esaminare un attimo il risultato a cui siamo giunti. Per l'uguaglianza fondamentale della trigonometria, sappiamo che:

$$\sin^2 x + \cos^2 x = 1$$

E quindi

$$\cos x = \pm\sqrt{1-\sin^2 x}$$

Potremmo applicare la stessa disuguaglianza nella nostra frazione, per trovarci nella condizione di dover calcolare il seno dell'arcoseno, operazione molto facile. Tuttavia, come si risolve il segno, che nella formula generale è doppio? Basta osservare che l'arcoseno da come risultati valori compresi tra $-\pi/2$ e $\pi/2$, e questo è proprio l'intervallo in cui il coseno è positivo – il segno sarà quindi $+$:

$$\frac{1}{\sqrt{1-\sin^2(\arcsin x)}} = \frac{1}{\sqrt{1-x^2}}$$

Questa volta il caso in cui $(Df)(f^{-1}(x)) = 0 \Rightarrow \sqrt{1-x^2} = 0$ esiste, ed è facile verificare che si verifica quando $x = \pm 1$. Quindi, l'intervallo dove è definita la derivata dell'arcoseno è $(-1;1)$.

Esempio 6.5.11 *Calcolare*

$$D\arccos x$$

In questo caso saltiamo i calcoli, praticamente identici a quelli fatti per calcolare l'arcoseno, e ci limitiamo a dare i risultati. Il dominio dell'arcoseno è sempre $[-1;1]$, ma questa volta la sua immagine cambia, e risulta essere $[0;\pi]$. La sua derivata è

$$-\frac{1}{\sqrt{1-x^2}}$$

Ovvero l'opposto della derivata dell'arcoseno.

6.6 Schema riassuntivo per il calcolo delle derivate

Ora che abbiamo calcolato la derivata delle principali funzioni e le regole per il loro calcolo, possiamo scrivere una tabella riassuntiva:

6.7 Utilità della derivata prima

6.7.1 Massimi e minimi relativi

x_0 è un punto di massimo (o minimo) relativo, se esiste un intorno del punto nel quale $f(x_0)$ è maggiore (minore) o uguale a tutti gli altri valori della funzione in quell'intorno. Scritto simbolicamente:

$$\begin{gathered} x_0 \text{ è max.rel. di } f(x) \\ \Updownarrow \\ \exists\delta > 0 : \forall x \in [x-\delta; x+\delta] : f(x_0) \geq x \end{gathered}$$

Ebbene, la derivata prima permette di calcolare dove si trovano questi punti. Intanto, cominciamo coll'enunciare la regola generale:

Teorema 6.7.1 *Se x_0 è un massimo o minimo relativo di $f(x)$, ed $f(x)$ è derivabile in questo punto, allora $f'(x_0) = 0$.*

Tuttavia questa regola non ci permette di enunciare l'affermazione simmetrica, ovvero che dove vi è derivata nulla, si trova un massimo o minimo relativo, poiché questo *non* è vero!

$f(x)$	$f'(x)$
k	0
x	1
x^n	nx^{n-1}
e^x	e^x
a^x	$a^x \ln a$
$\sin x$	$\cos x$
$\cos x$	$-\sin x$
$\tan x$	$\dfrac{1}{\cos^2 x} = 1 + \tan^2 x$
$\sinh x$	$\cosh x$
$\cosh x$	$\sinh x$
$\tanh x$	$\dfrac{1}{\cosh^2 x} = 1 - \tanh^2 x$
$\ln x$	$\dfrac{1}{x}$
$\log_a x$	$\dfrac{1}{x \ln a}$
$\arcsin x$	$\dfrac{1}{\sqrt{1-x^2}}$
$\arccos x$	$-\dfrac{1}{\sqrt{1-x^2}}$
$\arctan x$	$\dfrac{1}{1+x^2}$
$\operatorname{arcsinh} x$	$\dfrac{1}{\sqrt{1+x^2}}$
$\operatorname{arccosh} x$	$\dfrac{1}{\sqrt{x^2-1}}$
$\operatorname{arctanh} x$	$\dfrac{1}{1-x^2}$

Tabella 6.1: Derivate delle funzioni elementari

$[kf(x)]'$	$kf'(x)$
$f(x) \pm g(x)$	$f'(x) \pm g'(x)$
$[f(x)g(x)]'$	$f'(x)g(x) + f(x)g'(x)$
$\dfrac{1}{f(x)}$	$\dfrac{f'(x)}{f^2(x)}$
$\left[\dfrac{f(x)}{g(x)}\right]'$	$\dfrac{f'(x)g(x) - f(x)g'(x)}{g^2(x)}$
$\{f[g(x)]\}'$	$f(g)'g'(x)$
$(f^{-1})'(x)$	$\dfrac{1}{f'[f^{-1}(x)]}$

Tabella 6.2: Regole di derivazione

Esempio 6.7.1 *Studiare la derivata prima di*

$$f(x) = x^3$$

La sua derivata è:

$$f'(x) = 3x^2$$

Ora, la derivata si annulla in $x = 0$. Tuttavia, in questo punto non esiste né massimo né minimo relativo: infatti in un qualunque intorno di 0 esistono sia punti di f dal valore positivo sia negativo, mentre $f(0) = 0$. Dato quindi che nessun punto in cui $f'(x) = 0$ è massimo o minimo relativo, possiamo dire che la funzione non possiede né massimi né minimi.

La dimostrazione di questo teorema è la seguente.

Prendo un x_0, ad esempio minimo relativo della funzione, quindi, $\forall x \in [x_0 - \delta; x_0 + \delta] : f(x) \geq f(x_0)$. Se ora calcolo il rapporto incrementale tra uno qualunque di questi x ed x_0, ottengo il solito:

$$\frac{f(x) - f(x_0)}{x - x_0}$$

Questo rapporto ha sempre numeratore positivo o uguale a 0, mentre il denominatore, per un $x \in [x_0 - \delta; x_0]$, ha segno negativo o nullo, mentre per un $x \in [x_0; x_0 + \delta]$, avrà segno positivo o nullo. Quindi, il limite sinistro di tale funzione nel punto x_0 dovrà avere segno negativo o nullo, mentre a destra l'avrà positivo o nullo. Dato che la derivata, per ipotesi, nel punto x_0 esiste, dovrà essere nulla.

6.7.2 Relazione tra monotonia e segno di $f'(x)$

La relazione tra questi due fatti è spiegata dal seguente teorema:

Teorema 6.7.2 *Data una $f : I \to \mathbb{R}$, dove I è un intervallo, ed f derivabile in I, si ha che sono equivalenti i due seguenti:*

- *f è crescente su I*
- *$\forall x \in I : f'(x) \geq 0$ (la derivata della funzione su I è sempre positiva)*

Ovviamente, nel caso opposto si avrà una funzione decrescente ed una derivata sempre negativa.

Per dimostrare il teorema, bisogna prima dimostrare che la prima proposizione implica la seconda, quindi che la seconda implica la prima.

Per quanto riguarda la prima, si tratta di dimostrare che, a partire dall'ipotesi che:

$$\forall x_1, x_2 \in I, x_1 < x_2 : f(x_1) \leq f(x_2)$$

Si ha che

$$\forall x \in I : f'(x) \geq 0$$

Per fare ciò, prendo un $x \in I$, calcolo il rapporto incrementale tra questo x ed x_0:

$$\frac{f(x) - f(x_0)}{x - x_0}$$

Se $x_0 > x$, avrò che $x - x_0 \leq 0$ e che $f(x_0) \geq f(x)$ (per ipotesi la funzione è crescente ...), e quindi $f(x) - f(x_0) \leq 0$. Il rapporto incrementale risulta quindi il rapporto tra due valori negativi o nulli, e quindi il risultato sarà positivo o nullo. Nel caso opposto, ovvero $x_0 < x$, avrei invece il rapporto di due valori entrambi *positivi* o nulli, e quindi avrei sempre un rapporto positivo o nullo. Il limite per $x \to x_0$ deve quindi sempre darmi un risultato positivo o nullo, ma questo limite è la derivata della funzione, ovvero:

$$\forall x \in I : f'(x) \geq 0$$

Ed ecco che la prima parte del teorema è dimostrata

Per la seconda parte invece occorre utilizzare il cosiddetto *Teorema di Lagrange*, che dimostreremo poco più avanti (6.7.4)e che dice che, per una funzione f derivabile in $[a, b]$ e derivabile in (a, b), esiste sempre un punto c, interno all'intervallo, tale che

$$f'(x) = \frac{f(b) - f(a)}{b - a}$$

Quindi, tornando al nostra teorema, prendiamo due punti x_1 ed x_2: Lagrange dice che

$$\exists c \in (x_1; x_2) : f'(c) = \frac{f(x_1) - f(x_2)}{x_2 - x_1}$$

Ma $f'(c) \geq 0$ per ipotesi, $x_2 > x_2$ per costruzione, e quindi anche il numeratore della nostra frazione deve essere positivo o nullo:

$$f(x_2) - f(x_1) \geq 0$$

Ovvero

$$f(x_2) \geq f(x_1)$$

Ecco dimostrato il nostro teorema.

A questo punto possiamo dire qualcosa di più sulla ricerca dei massimi e dei minimi. Se prendiamo un punto x_0 nel quale la derivata cambia segno, ad esempio, da negativo (nell'intervallo $[x_0 - \delta; x_0]$) a positivo (nell'intervallo $[x_0; x_0 + \varepsilon]$, avremo che, per continuità, nel punto x_0 la funzione è nulla, e quindi potrebbe presentare un massimo o minimo. Ma, in effetti, nell'intervallo $[x_0 - \delta; x_0]$ la funzione è descrescente, dato che la derivata ha segno negativo, mentre in $[x_0; x_0 + \varepsilon]$ è crescente, e quindi abbiamo che:

$$\forall x \in [x_0 - \delta; x_0] : f(x) \geq f(x_0)$$

Dato che $x < x_0$, ed inoltre:

$$\forall x \in [x_0; x_0 + \varepsilon] : f(x) \geq f(x_0)$$

Poiché $x > x_0$. Unendo le due affermazioni, si ha che

$$\forall x \in [x_0 - \delta; x_0 + \varepsilon] : f(x) \geq f(x_0)$$

Dato quindi che in un intorno di x_0 la funzione è maggiore di $f(x_0)$, x_0 è un punto di minimo relativo. Il discorso inverso vale per punti di massimo relativo, e si ha una regola generale che ci dice che:

Teorema 6.7.3 *Se nel punto x_0 la funzione f' cambia di segno da negativo a positivo, si ha un punto di minimo relativo, mentre se cambia segno da positivo a negativo si ha invece un punto di massimo relativo.*

Quindi, se la funzione non cambia di segno, non si ha né massimo né minimo.

Esempio 6.7.2 *Studiare*

$$f(x) = x^2 e^{-x}$$

Calcoliamo la derivata prima:

$$f'(x) = 2xe^{-x} - x^2 e^{-x} = xe^{-x}(2 - x)$$

Questa funzione si annulla nei punto $x = 0$ ed $x = 2$. Nel primo punto si ha ovviamente un minimo, poiché in 0 la funzione vale 0, mentre per tutti gli altri punti ha sempre valore positivo. Nel punto 2 invece? Se esaminiamo il segno della derivata prima, otteniamo che nell'intervallo $(-\infty, 0)$ è negativa, in $(0; 2)$ è positiva e infine in $(2; +\infty)$ è di nuovo negativa. Quindi, nel punto 0 si ha un minimo come già avevamo constatato, mentre nel punto 2 si ha un massimo.

Ora che abbiamo trovato a grandi linee il comportamento della funzione in un intervallo finito, vediamo il suo comportamento $\pm\infty$:

$$\lim_{x \to +\infty} f(x) = \lim_{x \to +\infty} \frac{x^2}{e^x} = 0$$

e

$$\lim_{x \to -\infty} f(x) = \lim_{x \to -\infty} \frac{x^2}{e^x} = +\infty$$

Quindi il grafico della funzione avrà approssimativamente questo andamento:

Dato che è una funzione continua che ha per dominio un intervallo ($\mathbb{R}$), avrà un intervallo anche per sua immagine, e dato che contiene i valori 0 come minimo e $+\infty$ come massimmo, tale intervallo sarà $\mathbb{R}^+$, per il teorema di Bolzano.

Esempio 6.7.3 *Per quali valori di λ, l'equazione*

$$x^2 e^{-x} = \lambda$$

Ha 3 soluzioni distinte?

Noi sappiamo che il problema di trovare le soluzioni di un'equazione $f(x) = k$ può anche essere visto in chiave geometrica, ovvero diventa il problema di trovare le intersezioni tra la retta $y = k$ e la curva $y = f(x)$. Nel nostro caso abbiamo già il grafico della curva $y = f(x)$, e ci viene chiesto in pratica quali rette $y = \lambda$ intersecano in 3 punti questo grafico. Esaminiamo i vari casi che si presentano:

- *Per $\lambda < 0$, non ci sono intersezioni.*
- *Per $\lambda = 0$, è presente una sola intersezione ($x = 0$).*
- *Per $0 < \lambda < f(2) = 4/e^2$, si hanno tre intersezioni, ($x_1 < 0,\ 0 < x_2 < 2,\ x_3 > 2$).*
- *Per $\lambda = 4/e^2$, si hanno due intersezioni ($x_1 < 0,\ x_2 = 2$).*
- *Per $\lambda > 4/e^2$, si ha una sola intersezione ($x < 0$).*

Come si può notare, dal semplice studio del grafico della funzione si è ricavata una notevole mole di dati. La soluzione del problema sarà quindi:

$$0 < \lambda < \frac{4}{e^2}$$

Da notare anche il fatto che noi non abbiamo per nulla risolto l'equazione, cosa d'altronde molto probabilmente impossibile con gli strumenti a nostra disposizione.

Un teorema che avevamo lasciato non dimostrato era:

Teorema 6.7.4 (Teorema di Lagrange, o del valor medio) *Presa una funzione f:*

$$f : [a; b] \to \mathbb{R}$$

continua su $[a; b]$ e derivabile su $(a; b)$, allora:

$$\exists c \in (a; b) : f'(c) = \frac{f(b) - f(a)}{b - a}$$

Ovvero esiste sempre almeno un punto interno all'intervallo nel quale la tangente alla funzione è parallela alla retta passante per i punti agli estremi dell'intervallo.

Inanzitutto, spieghiamo l'equivalenza tra la formula data e la sua spiegazione. Sappiamo che $f'(c)$ è il coefficiente angolare della retta tangente alla funzione f nel punto c. Il secondo membro dell'uguaglianza in realtà non è altro che l'espressione per trovare il coefficiente angolare della retta passante per i punti $(a, f(a))$ e $(b, f(b))$, che sono per l'appunto i punti estremi dell'intervallo. Se queste due rette hanno coefficiente angolare uguale, significa che hanno la stessa pendenza, ovvero che sono parallele.

Da notare subito che questo teorema non esclude che i punti in cui ciò avviene siano addirittura più di uno!

Per dimostrarlo si parte inanzitutto col caso in cui $f(a) = f(b)$; l'enunciato del teorema diventa quindi il seguente:

Teorema 6.7.5 (Teorema di Rolle) *Presa una funzione f:*

$$f : [a; b] \to \mathbb{R}$$

continua su $[a; b]$, derivabile su $(a; b)$ e con valori uguali agli estremi ($f(a) = f(b)$), allora:

$$\exists c \in (a; b) : f'(c) = 0$$

Ovvero esiste sempre almeno un punto interno all'intervallo nel quale la tangente alla funzione è parallela all'asse x.

Se la funzione $f(x)$ è costante, il teorema è dimostrato, in quanto non solo uno, ma *tutti* i punti nell'intervallo hanno derivata nulla.

In caso contrario, dato che la funzione è considerata *continua*, allora, per il teorema di Weierstrass, possiede minimo (m) e massimo (M) assoluti. Sicuramente, $m < M$, poiché la funzione non è costante. Quindi, abbiamo che:

$$\begin{aligned} &m = f(\alpha) \\ &M = f(\beta) \\ &m < M \implies \alpha \neq \beta \end{aligned}$$

Dato che $f(a) = f(b)$, sicuramente non può essere che $a = \alpha$ *e* $b = \beta$, altrimenti avremmo $f(\alpha) = f(\beta)$, ovvero $m = M$. Quindi, almeno uno dei due punti deve essere interno all'intervallo, ma in quel caso, avremo che o il massimo, o il minimo, è interno all'intervallo, e quindi in questi punti la funzione avrà derivata nulla. CVD.

Passando nuovamente al caso generale, come vedremo potremo servirci di questo risultato parziale per costruire la nostra dimostrazione. Adesso siamo di nuovo nel caso in cui $f(a)$ ed $f(b)$ non devono necessariamente essere uguali, e cerchiamo un punto in cui

$$f'(c) = \frac{f(b) - f(a)}{b - a}$$

Prendiamo ora una funzione ausiliaria, chiamiamola F, così definita:

$$F(x) = f(x) - \frac{f(b) - f(a)}{b - a}(x - a)$$

Dato che abbiamo sottratto ad f (derivabile) un polinomio di primo grado in x (derivabile anch'esso), avremo quindi ancora una funzione derivabile, ovvero la nostra F. A questo punto calcoliamo il valore di questa nuova funzione in a e b:

$$F(a) = f(a) - \frac{f(b) - f(a)}{b - a}(a - a) = f(a)$$

$$F(b) = f(b) - \frac{f(b) - f(a)}{b - a}(b - a) = f(b) - [f(b) - f(a)] = f(a)$$

Quindi, F soddisfa le ipotesi del teorema di Rolle, dato che è continua, derivabile ed ha uguali valori agli estremi. Quindi, esiste almeno un punto $c \in (a;b)$ tale per cui vale:

$$F'(c) = 0$$

Ma

$$F'(x) = f'(x) - \frac{f(b) - f(a)}{b - a}$$

E quindi la relazione trovata equivale a:

$$f'(c) - \frac{f(b) - f(a)}{b - a} = 0$$

Che è poi:

$$f'(c) = \frac{f(b) - f(a)}{b - a}$$

Ecco trovato il nostro punto c. CVD.

6.7.3 La regola di De L'Hospital

La regola di De L'Hospital è un teorema utile nel calcolo di alcune forme di indeterminazione.

Criterio 6.7.6 (Regola di De L'Hospital) *Se nel calcolo di un limite*

$$\lim_{x \to c} \frac{f(x)}{g(x)}$$

dove c è un valore finito o infinito, ci si trova in una forma di indeterminazione del tipo

$$\frac{0}{0}$$

O

$$\frac{\infty}{\infty}$$

Allora, se

$$\lim_{x \to c} \frac{f'(x)}{g'(x)} = l$$

anche il rapporto di partenza ammette limite, e tale limite è proprio l.

Una importante osservazione da fare subito è che, nel caso in cui $f'(x)/g'(x)$ non abbia limite, *questo non significa che il limite di partenza non esista!* La regola di De L'Hospital, come altri criteri visti in precedenza, vale solo nel caso siano soddisfatte certe ipotesi, in caso contrario non da alcuna indicazione a proposito.

Esempio 6.7.4 *Calcolare*

$$\lim_{x \to 0} \frac{\sin x - x}{x^3}$$

Ci troviamo in una forma di indeterminazione, del tipo $0/0$. Utilizzando De L'Hospital, otteniamo:

$$\lim_{x \to 0} \frac{\cos x - 1}{3x^2}$$

Ricordando che

$$\lim_{x \to 0} \frac{\cos x - 1}{x^2} = \frac{1}{2}$$

Otteniamo

$$\lim_{x \to 0} \frac{\cos x - 1}{3x^2} = -\frac{1}{6}$$

E dato che il rapporto preso in considerazione ammette limite, allora anche il rapporto originario lo ammette, e vale proprio $-1/6$.

Esempio 6.7.5 *Calcolare*

$$\lim_{x \to 0} \frac{\arctan x - x}{x^3}$$

Ci troviamo nella stessa condizione di prima: quindi, come prima, applichiamo De L'Hospital:

$$\lim_{x \to 0} \frac{\frac{1}{1+x^2} - 1}{3x^2} = \lim_{x \to 0} \frac{-\frac{x^2}{1+x^2}}{3x^2} = \lim_{x \to 0} -\frac{1}{3} \cdot \frac{1}{1+x^2} = -\frac{1}{3}$$

6.8 Derivate di ordine superiore

Abbiamo già definito la derivata di primo ordine; ora defininiamo la derivata di secondo ordine come derivata della derivata del primo ordine, ovvero:

$$f''(x) = (f')'(x)$$

In generale, per definire la derivata di ordine n-simo di una funzione f si usa questa definizione "ricorsiva":

$$f^n x = (f^{n-1})'(x)$$

6.8.1 Massimi e minimi relativi con la derivata seconda

Prendiamo una funzione f, derivabile due volte in un intorno del punto x_0. Allora abbiamo la seguente regola:

Criterio 6.8.1 *Se $f'(x_0) = 0$ ed $f''(x_0) \neq 0$, allora se $f''(x_0) > 0$, la funzione ha in x_0 un minimo relativo, altrimenti ($f''(x_0) < 0$) la funzione avrà un massimo relativo.*

Come sempre, nei casi non considerati ($f'(x_0) = f''(x_0) = 0$), nulla si può dire.

La spiegazione della regola è la seguente. Prendiamo ad esempio il caso in cui $f''(x_0) > 0$: avremo che

$$f''(x_0) = \lim_{x \to x_0} \frac{f'(x) - f'(x_0)}{x - x_0} = \lim_{x \to x_0} \frac{f'(x)}{x - x_0}$$

Ora, il rapporto incrementale dovrà essere sempre maggiore di zero in un intorno di x_0. Nella parte sinistra dell'intevallo si avrà che $x - x_0 < 0$, e quindi $f'(x) < 0$, mentre nella parte destra si avrà il caso contrario, ovvero $f'(x) > 0$. Quindi nel punto x_0 la derivata cambia segn da negativa a positiva, e quindi si ha un minimo relativo. CVD.

Come già detto, nel caso in cui $f''(x_0) = 0$, non si può dire nulla, perchè in un intorno di $f''(x_0)$ in tal caso si possono trovare sia valori positivi che negativi della derivata prima.

Esempio 6.8.1 *Considerare la validità o meno del criterio appena esposto per le seguenti funzioni:*

$$x^3, \quad in\ x_0 = 0$$

$$x^4, \quad in\ x_0 = 0$$

$$\sin x, \quad in\ x_0 = \frac{\pi}{2}$$

Per $f(x) = x^3$ si ha:

$$f'(x) = 3x^2; \quad f'(x_0) = f'(0) = 0$$

E

$$f''(x_0) = 6x; \quad f''(x_0) = f''(0) = 0$$

In questo caso non si può dire nulla, e, se si esamina la funzione, si nota che in 0 non esiste né minimo né massimo relativo.

Per $f(x) = x^4$, si ha:

$$f'(x) = 4x^3; \quad f'(x_0) = f'(0) = 0$$

E

$$f''(x) = 14x^2; \quad f''(x_0) = f''(0) = 0$$

Anche qui quindi non possiamo dire nulla su massimi o minimi della funzione, che questa volta invece risulta avere in 0 un minimo.

Infine, per $f(x) = \sin x$, si ha:

$$f'(x) = \cos x; \quad f'(x_0) = f'(\pi/2) = 0$$

E

$$f''(x) = -\sin x; \quad f''(x_0) = f''(\pi/2) = -1$$

Quindi, la funzione in $\pi/2$ presenta un massimo relativo.

6.8.2 Concavità, convessità e flessi

Per quanto riguarda la derivata prima, il suo segno era collegato con la monotonia della funzione – attraverso lo studio del segno della derivata seconda invece cosa possiamo ricavare?

Ad esempio, se troviamo che $f''(x) > 0$, per $x \in I$, possiamo intanto dire che $f'(x)$ è *crescente* in tale intervallo. Questo significa che il grafico della funzione in I ha pendenze sempre maggiori man mano che si prendono x più grandi. Disegnando una funzione con questa caratteristica, si nota che possiede la particolarità che il grafico all'interno di I sta sempre al di sopra di ogni tangente al grafico stesso – la definizione che abbiamo appena dato è quella di *convessità*.

In termini formali, una funzione f è convessa se:

$$\forall x, x_0 \in I : f(x) \geq f(x_0) + f'(x_0)(x - x_0)$$

Infatti, $f(x_0) + f'(x_0)(x - x_0)$ rappresenta la retta tangente in x_0: la proposizione data dice quindi che, per qualunque x che prendiamo dell'intervallo, tale punto sarà sopra alla tangente ad un qualunque punto della funzione nell'intervallo.

Le osservazioni fatte ricadono tutte nel seguente teorema:

Teorema 6.8.2 *Se $f : I \to \mathbb{R}$ è derivabile due volte su I, allora sono equivalenti le seguenti due affermazioni:*

- $\forall x \in I : f''(x) \geq 0$
- *f è convessa in I*

Nel caso opposto, ovvero se la derivata seconda è sempre negativa, si ha che la funzione è *concava*, che è ovviamente il concetto opposto (la funzione sta sempre sotto le sue derivate).

Per questo teorema, si possono trovare alcune funzioni convesse: ad esempio, qualunque potenza pari:

$$f(x) = x^{2n}, n \in \mathbb{N}$$

è convessa, poiché:

$$f''(x) = (2n \cdot x^{2n-1})' = 2n(2n-1)x^{2n-2}$$

Che è sempre positiva.

Inoltre anche una qualunque esponenziale è convessa, infatti:

$$(a^x)'' = (a^x \ln a)' = a^x \ln^2 a$$

Che è il prodotto di due quantità sempre positive.

Per logaritmo si ha invece:

$$(\log_a x)'' = \left(\frac{1}{x} \cdot \frac{1}{\ln a}\right)' = \left(-\frac{1}{x^2} \cdot \frac{1}{\ln a}\right)$$

Che è sempre negativo – la funzione logaritmo è quindi invece sempre concava.

Si può aggiungere a queste definizioni anche quella di *flesso*: si dice che f ha un flesso in x_0 se $f''(x_0) = 0$ ed f'' cambia segno attraversando x_0. Quindi, si ha un flesso se la funzione cambia da concava a convessa o viceversa.

Inoltre, nel caso in cui $f'(x_0) = 0$ si dice che il flesso è *a tangente orizzontale*. Osserviamo quindi che, quando fallisce il criterio per trovare massimi e minimi relativi per mezzo della derivata seconda è possibile che ci si trovi in questo caso, ma perché ciò avvenga è necessaria anche una condizione in più (il cambio di segno della derivata seconda).

6.8.3 Metodo di Newton

Il metodo di Newton è un metodo per ricercare le soluzioni di un'equazione nella forma $f(x) = 0$ – è un sistema molto potente, tanto che la maggior parte degli altri metodi che hanno lo stesso scopo si basano su questo.

Tale metodo è applicabile se sono soddisfatte le seguenti condizioni:

1. $f : I \to \mathbb{R}$ è derivabile due volte su I.
2. $f(a) > 0, f(b) < 0$
3. $f'(x) < 0, \forall x \in I$ (oppure si possono prendere i contrari di queste ultime due condizioni).
4. $f''(x) \geq 0$ (o ≤ 0).

Come si nota, pur essendo, come vedremo, un metodo più potente di altri, richiede anche più informazioni.

Dalla seconda condizione sappiamo che, per il teorema degli zeri, esiste almeno una soluzione in (a, b). Dalla terza condizione inoltre sappiamo che la soluzione è anche unica, poiché la funzione è decrescente. Quindi:

$$\exists! c \in I : f(c) = 0$$

L'esecuzione del metodo di Newton è la seguente: si parte prendendo un x_0 tale che:

$$a \leq x_0 < \alpha$$

(d'ora in poi chiameremo α la soluzione dell'equazione). Sicuramente, un punto che soddisfa tale uguaglianza è a stesso. Se ora tracciamo la tangente al grafico nel punto x_0, questa incrocerà l'asse delle x in una ben precisa ascissa, che chiameremo x_1, e che si ricava risolvendo il seguente sistema:

$$\begin{cases} y = 0 \\ y = f(x_0) + f'(x_0)(x_1 - x_0) \end{cases}$$

$$x_1 = \frac{f'(x_0)x_0 + f(x_0)}{f'(x_0)}$$

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$$

La divisione per $f'(x_0)$ è consentita poiché tale valore è, per ipotesi, minore di 0. Inoltre, dato che $a \leq x_0 < \alpha$, e $f(a) > 0$ e $f(\alpha) = 0$, avremo che $f(x_0) > 0$. Inoltre, come già detto, $f'(x_0) < 0$, e quindi:

$$\frac{f(x_0)}{f'(x_0)} < 0$$

E quindi, sottratto ad x_0, danno un risultato maggiore di x_0. Quindi:

$$x_0 < x_1$$

x_1 è quindi più vicino di x_0 alla soluzione α, ma solo a condizione che sia ancora minore di α stesso. Per dimostrare ciò basterà mostrare che $f(x_1) > 0$ (per il teorema degli zeri). Dato che f è convessa per ipotesi, allora avremo che:

$$f(x_1) > f(x_0) + f'(x_0)(x_1 - x_0)$$

Non abbiamo fatto altro che prendere la definizione di funzione convessa e sostituire ai due generici x, x_0 i valori x_0 ed x_1 (l'omonimia tra le due variabili nella definizione e in questa dimostrazioni non confonda!). Ma il secondo membro dell'uguaglianza rappresenta l'ordinata della retta tangente in x_0 nel punto x_1, che è stato per l'appunto definito dicendo che è il punto dove la retta tangente ad x_0 incrocia l'asse delle ascisse – quindi, vale 0. La disuguaglianza diventa quindi:

$$f(x_1) \geq 0$$

e quindi

$$x_1 \leq \alpha$$

Si può dimostrare poi che $x_1 \neq \alpha$, e quindi che:

$$x_1 < \alpha$$

A questo punto, si ha quindi che:

$$x_0 < x_1 < \alpha$$

Ora, se osserviamo x_1, notiamo che esso soddisfa le condizioni di partenza che abbiamo usato per prendere x_0, e quindi possiamo prendere questo punto per partire di nuovo con gli stessi calcoli, ottenendo quindi un x_2 tale che:

$$x_0 < x_1 < x_2 < \alpha$$

E così via. Quella che andiamo a formare è, quindi, una successione così definita:

$$\begin{cases} x_0 \in [0, \alpha) \\ x_{n+1} = x_n - \dfrac{f(x_n)}{f'(x_n)} \end{cases}$$

Questa successione è chiaramente crescente e limitata, e quindi ammette un limite finito, che sarà d'altronde $\widetilde{\alpha} \leq \alpha$ (con $\widetilde{\alpha}$ indicheremo per l'appunto il limite di tale successione). Ma in realtà questo limite vale proprio α! Infatti, se prendiamo la seconda uguaglianza che definisce la serie, e calcoliamo il limite di entrambi i membri, per $n \to +\infty$, avremo ancora due quantità uguali:

$$\lim_{n\to+\infty} x_{n+1} = \lim_{n\to+\infty} x_n = \widetilde{\alpha}$$

$$\lim_{n\to+\infty} x_n = \widetilde{\alpha}$$

Ricordando poi che f ed f' sono continue:

$$\lim_{n\to+\infty} f(x_n) = f(\widetilde{\alpha})$$

$$\lim_{n\to+\infty} f'(x_n) = f(\widetilde{\alpha})$$

Quindi, sostituendo nell'uguaglianza citata sopra, otteniamo:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

$$\Downarrow$$

$$\widetilde{\alpha} = \widetilde{\alpha} - \frac{f(\widetilde{\alpha})}{f'(\widetilde{\alpha})}$$

Che, semplificando, equivale a:

$$\frac{f(\widetilde{\alpha})}{f'(\widetilde{\alpha})} = 0$$

Ma, dato che, per ipotesi, f' è sempre negativo (e quindi diverso da zero), possiamo semplificarlo:

$$f(\widetilde{\alpha}) = 0$$

Ma all'inizio della nostra dimostrazione avevamo detto che l'unico punto in cui la funzione si annulla è il nostro α, quindi $\widetilde{\alpha} = \alpha$:

$$\alpha = \lim_{n \to +\infty} x_n$$

Per il caso in cui la derivata seconda sia negativa, l'unico cambiamento sta nella scelta dell'x_0:

$$x_0 \in (\alpha, b]$$

Esempio 6.8.2 *Calcolare un valore approssimato per* $\sqrt{131}$.

Noi sappiamo che $\sqrt{131}$ *è anche la soluzione dell'equazione:*

$$131 - x^2 = 0$$

Quindi verifichiamo se possiamo usare il metodo di Newton per $f(x) = 131 - x^2$. *Prendiamo inanzitutto i due estremi* a *e* b, *che potrebbero ad esempio essere* $a = 10$ *e* $b = 12$:

$$f(a) = f(10) = 131 - 10^2 = 131 - 100 = 31 > 0$$

$$f(b) = f(12) = 131 - 12^2 = 131 - 144 = -11 < 0$$

La derivata f' *è:*

$$f'(x) = -2x$$

Che nell'intervallo $[a; b]$ *è effettivamente sempre negativa. Infine:*

$$f''(x) = -2$$

Che ovviamente è sempre negativa anch'essa. Possiamo quindi applicare il metodo di Newton. Il valore x_{n+1} *varrà quindi:*

$$x_n - \frac{f(x_n)}{f'(x_n)} = x_n + \frac{131 - x_n^2}{2x_n} = \frac{x_n^2 + 131}{2x_n} = \frac{1}{2}\left(x_n + \frac{131}{x_n}\right)$$

Quindi il valore x_{n+1} *è il punto medio tra* x_n *e* $131/x_n$. *La successione risultante sarà quindi:*

$$\begin{cases} x_0 = 12 \\ x_{n+1} = \dfrac{1}{2}\left(x_n + \dfrac{131}{x_n}\right) \end{cases}$$

e $\sqrt{131} = \lim_{n \to \infty} x_n$.

Esempio 6.8.3 *In alcune applicazioni dell'ottica si utilizza l'approssimazione*

$$x = \tan x$$

Determinare approssimativamente dove questo è vero.

In pratica si tratta di determinare la soluzione dell'equazione data: se si prende la soluzione solo nell'intervallo $[0; 2\pi)$ *e si esclude la soluzione banale* $x = 0$, *l'unica altra soluzione che si presenta è compresa tra* π *e* $3\pi/2$.

Inanzitutto osserviamo che l'equazione proposta è equivalente a:

$$\sin x - x \cos x = 0$$

E che quindi non è risolvibile elementarmente: proviamo ad usare il metodo di Newton.

Come primo intervallo che viene in mente da usare c'è proprio $[\pi; 3\pi/2]$, *il quale va bene per la seconda condizione del metodo:*

$$f(\pi) = \sin \pi - \pi \cos \pi = 0 - \pi(-1) = \pi$$

$$f(\frac{3}{2}\pi) = sin\frac{3}{2}\pi - \frac{3}{2}\pi \cos \frac{3}{2}\pi = -1 - \frac{3}{2}\pi \cdot 0 = -1$$

Ma che non soddisfa la terza condizione, infatti:

$$f'(x) = \cos x - (\cos x - x \sin x) = x \sin x$$

$$f'(a) = \pi \sin \pi = 0$$

Tuttavia, sostituendo a π un valore leggermente maggiore, come $5\pi/4$ (valore medio tra i vecchi a e b), la situazione va di nuovo bene per le prime due condizioni:

$$f(a) = f(\frac{5}{4}\pi) = sin\frac{5}{4}\pi - \frac{5}{4}\pi \cos\frac{5}{4}\pi = -\frac{\sqrt{2}}{2} + \frac{5}{4}\pi \cdot \frac{\sqrt{2}}{2} = \frac{\sqrt{2}}{8}(5\pi - 4) > 0$$

Poichè $\pi > 3$, e quindi $5\pi > 15$ e $5\pi - 4 > 11 > 0$. Ora, studiamo la derivata prima:

$$x \sin x$$

x, ovviamente, per $x \in [5\pi/4; 3\pi/2]$, è sempre positiva. $\sin x$ è compreso invece tra $\sin(5\pi/4)$ e $\sin(3\pi/2)$ (si può arrivare facilmente a questo risultato combinando il teorema di Bolzano con il fatto che in questo intervallo il seno è strettamente monotona – decrescente, per la precisione), ovvero tra $-\sqrt{2}/2$ e -1. Quindi, in ogni punto di I, il prodotto $x \sin x$ è un valore negativo, e quindi anche la condizione necessaria per la derivata prima è soddisfatta.

Per quanto riguarda la derivata seconda:

$$f''(x) = \sin x + x \cos x$$

Si ha che è sempre minore di zero. Infatti la condizione:

$$\sin x + x \cos x < 0$$

Equivale a:

$$\sin x < -x \cos x$$

Ovvero (il coseno è negativo in questo intervallo!):

$$\frac{\sin x}{\cos x} > -x$$
$$\tan x > -x$$

Ed effettivamente, in $[5\pi/4, 3\pi/2]$, la tangente assume valori positivi, mentre $-x$ sarà ovviamente negativa, quindi anche questa condizione è soddisfatta.

Ora possiamo definire la successione. Per x_0 prendiamo la a stessa:

$$x_0 = \frac{5}{4}\pi$$

x_{n+1} avrà questa forma:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{\sin x_n - x_n \cos x_n}{x_n \sin x_n} = x_n - 1 + \cot x_n$$

Riassumendo il tutto, otteniamo che il valore cercato equivale al limite della seguente successione:

$$\begin{cases} x_0 = \frac{5}{4}\pi \\ x_{n+1} = x_n - 1 + \cot x_n \end{cases}$$

Dopo sole 5 iterazioni, già si arriva al risultato approssimato

$$\lim_{x \to \infty} x_n \approx 4.493409457909064$$

(utilizzando per π un'approssimazione ad 11 cifre), con una precisione di 10^{-15}: si può quindi notare come, dopo un bassissimo numero di iterazioni il risultato abbia già acquisito una precisione veramente notevole

Detto per inciso, il risultato trovato è circa 1.43π, quindi molto vicino a 1.5π, ovvero $3\pi/2$, che era l'estremo superiore dell'intervallo. Il risultato del problema è quindi che l'approssimazione è valida per valori molto vicini a $3\pi/2$ o, ancora meglio, ad 1.43π.

6.9 Linearizzazione e differenziale

Un'operazione spesso utile è quella di approssimare una funzione ad una retta, per avere così una semplice dipendenza lineare rispetto alla variabile indipendente, molto facile da manipolare.

Un esempio di questa linearizzazione è semplicemente quello di approssimare l'incremento che subisce la funzione quando la variabile indipendente varia da x_0 a $x_0 + \mathrm{d}x$.

Più precisamente, sia $f : (a, b) \to \mathbb{R}$ derivabile in $x_0 \in (a, b)$, ed un valore $\mathrm{d}x$ (che normalmente si pensa in valore assoluto molto piccolo). Allora, a questo punto approssimiamo la funzione con la retta tangente ad essa nel punto x_0. L'incremento che subisce questa approssimazione passando da $x = x_0$ ad $x = x_0 + \mathrm{d}x$ è pari a

$$\tan \alpha \cdot \mathrm{d}x$$

Dove α è l'angolo che la retta forma con l'asse delle ascisse. Ma $\tan\alpha = f'(x_0)$, e quindi abbiamo:

$$f'(x_0)\,\mathrm{d}x$$

Questa variazione viene chiamata *differenziale di f nel punto* x_0 e si indica con il simbolo $\mathrm{d}f(x_0)$:

$$\mathrm{d}f(x_0) = f'(x_0)\,\mathrm{d}x$$

A questo punto cerchiamo l'errore commesso nell'approssimazione, o per lo meno definiamolo qualitativamente se non quantitativamente. L'errore sarà:

$$[f(x_0 + \mathrm{d}x) - f(x_0)] - \mathrm{d}f(x_0) = \Delta f - \mathrm{d}f(x_0)$$

Partiamo dalla definizione di derivata di f in x_0:

$$\lim_{\mathrm{d}x\to 0} \frac{f(x_0 + \mathrm{d}x) - f(x_0)}{\mathrm{d}x} = f'(x_0)$$

Che diventa, utilizzando la scrittura fuori dal segno di limite:

$$\frac{f(x_0 + \mathrm{d}x) - f(x_0)}{\mathrm{d}x} = f'(x_0) + \varepsilon(\mathrm{d}x), \quad \lim_{\mathrm{d}x\to 0} \varepsilon(\mathrm{d}x) = 0$$

Ovvero, dove $\varepsilon(\mathrm{d}x)$ è un *infinitesimo* per $\mathrm{d}x \to 0$. Moltiplicando per $\mathrm{d}x$, otteniamo:

$$f(x_0 + \mathrm{d}x) - f(x_0) = f'(x_0)\,\mathrm{d}x + \mathrm{d}x \cdot \varepsilon(\mathrm{d}x)$$

Ovvero

$$\Delta f - \mathrm{d}f(x_0) = \mathrm{d}x \cdot \varepsilon(\mathrm{d}x)$$

Quindi l'errore commesso è pari a $\mathrm{d}x$ per una quantità che, per $\mathrm{d}x \to 0$, tende a zero, che dunque è un *infinitesimo di ordine superiore a* $\mathrm{d}x$.

A questo punto, introduciamo la seguente notazione per scrivere pi chiaramente la relazione appena trovata. Scrivendo:

$$f = o(g), \text{ per } x \to c$$

Che si legge proprio *o piccolo di g* si indica una funzione tale che:

$$\lim_{x\to c} \frac{f(x)}{g(x)} = 0$$

Ovvero, una funzione che tende a 0 *più velocemente* di g, per x che tende a c (nel caso di infinitesimi).

Ad esempio

$$f = o(1)$$

Significa semplicemente che

$$f(x) \to 0$$

Oppure

$$f = o(x)$$

Significa che

$$\frac{f(x)}{x} \to 0$$

Sempre intendendo che per entrambe le stesse espressioni, x tenda allo stesso c, finito o infinito.

Con questa scrittura, quindi, otteniamo che:

$$\Delta f - \mathrm{d}f(x_0) = o(\mathrm{d}x), \text{ per } \mathrm{d}x \to 0$$

Ovvero: l'errore che si commette a sostituire a Δf il differenziale della funzione è un o piccolo di $\mathrm{d}x$. A volte si scrive anche, più sinteticamente:

$$\Delta f \approx \mathrm{d}f \text{ al prim'ordine, vicino a } x_0$$

6.10 Formula di Taylor

Lo scopo della formula di Taylor è quello di approssimare una qualunque funzione per mezzo di un polinomio – questo è molto utile, poiché i polinomi sono le funzioni più semplici da manipolare.

Inanzitutto, è ovvio che non è possibile scrivere una funzione che non sia un polinomio stesso sotto forma di poliniomio, altrimenti sarebbero già polinomi queste funzioni; quel che si può sperare è di avere un'approssimazione la migliore possibile. Per fare ciò, si può dimostrare, si deve obbligatoriamente scegliere un punto nel quale l'approssimazione sarà perfetta e dal quale, man mano che ci si allontana, l'approssimazione peggiora. In generale, però, che significato bisogna attribuire al concetto di "approssimare"? Il significato più ovvio è quello di minimizzare l'errore che si compie sostituendo ad $f(x)$ il poliniomio trovato, chiamiamolo $P_n(x)$, dove la n indica l'ordine del polinomio considerato. Questo errore sarà espresso dalla formula:

$$\varepsilon(x) = f(x) - P_n(x)$$

Quindi, inanzitutto, potremmo chiedere che l'errore nel punto x_0 dove pretendiamo la migliore approssimazione sia, per l'appunto, nullo:

$$\varepsilon(x_0) = 0$$

Quindi, si potrebbe chiedere che la retta tangente ad $f(x)$ e $P_n(x)$ sia la stessa, e che quindi abbiano la derivata prima identica, ovvero che:

$$\varepsilon'(x_0) = 0$$

Poi, ancora, si potrebbe chiedere che la loro concavità o convessità sia la stessa, ovvero che abbiano la stessa derivata seconda in x_0, e cioé che:

$$\varepsilon''(x_0) = 0$$

Insomma, continuando con questo meccanismo, in pratica si tratta di porre a 0 tutte le successive derivate della funzione errore. Applicando n volte questo processo, si avranno n condizioni, che quindi poste a sistema permettono di calcolare il polinomio $P_n(x)$.

Allora, presa una $f : I \to \mathbb{R}$ che ammetta derivate di ogni ordine (ovvero che ammetta derivata prima, seconda, terza, ecc ...) ed un punto $x_0 \in I$, allora il polinomio cercato è il seguente:

$$\begin{aligned} P_n(x) = & \sum_{i=0}^{n} \frac{f^i(x_0)}{i!}(x - x_0)^i = \\ & f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \ldots + \frac{f^n(x_0)}{n!}(x - x_0)^n \end{aligned}$$

Facciamo ora una considerazione sull'errore che si compie, ovvero il valore che bisognerebbe aggiungere a $P_n(x)$ per avere esattamente la funzione cercata. Dimostriamo che la condizione posta per trovare il polinomio, ovvero che:

$$\varepsilon^i(x_0) = 0, i \leq n$$

Equivale a

$$f(x) - P_n(x) = o((x - x_0)^n), \text{ per } x \to x_0$$

Ovvero, che l'errore è un o piccolo di $(x - x_0)^n$. Per dimostrare questa equivalenza, occorre ovviamente prima dimostrarla in un verso e poi in quello opposto.

Inanzitutto, per semplicità, chiamiamo $\varepsilon(x)$ la differenza $f(x) - P_n(x)$, utilizzando quindi la solita notazione. Cerchiamo ora di dimostrare che

$$\varepsilon(x) = o((x - x_0)^n)$$

Ovvero che:

$$\lim_{x \to x_0} \frac{\varepsilon(x)}{(x - x_0)^n} = 0$$

Per fare ciò prendiamo un caso in particolare, ovvero per $n = 2$, ma che è facilmente estendibile a qualunque grado. Ci troviamo quindi nell'ipotesi che:

$$\varepsilon(x_0) = \varepsilon'(x_0) = \varepsilon''(x_0) = 0$$

E dobbiamo dimostrare che:

$$\frac{\varepsilon(x)}{(x - x_0)^2} \to 0, \quad x \to x_0$$

Utilizzando De L'Hospital (poichè ci troviamo in una forma di indeterminazione 0/0), abbiamo:

$$\frac{\varepsilon'(x)}{2(x - x_0)}$$

E ci troviamo ora di nuovo nella stessa condizione di prima; applichiamo quindi di nuovo De L'Hospital:

$$\frac{\varepsilon''(x)}{2}$$

Che effettivamente, per ipotesi, tende a 0. Quindi, ripercorrendo i vari limiti considerati, otteniamo che anche il primo tende a 0. In generale, se ci troviamo ad aver calcolato un polinomio di ordine n, bastera' applicare n volte De L'Hospital.

D'altronde, anche l'equivalenza letta in senso contrario è vera. Infatti, ragionando per assurdo, se $\varepsilon(x)/(x-x_0)^2$ non tendesse a zero, dovrebbe tendere a $\pm\infty$, in quanto il suo denominatore tende a zero. Invece, applicando De L'Hospital abbiamo:

$$\frac{\varepsilon'(x)}{2(x-x_0)}$$

Formula per la quale ci troviamo nella stessa condizione di prima. In questo caso e' sufficiente una sola altra applicazione del teorema di De L'Hospital, ed in generale ne occorrono n, per arrivare alla conclusione che il limite iniziale e' pari a $\varepsilon''(x)$ (o in generale $\varepsilon^n(x)$), che ovviamente non tende a $\pm\infty$.

Esempio 6.10.1 *Calcolare, nel punto $x_0 = 0$, il polinomio di Taylor di ordine generico n della funzione:*

$$f(x) = e^x$$

In questo caso abbiamo $f(x) = f'(x) = f''(x) = \cdots = f^n(x) = e^x$ e quindi $f^n(x_0) = 1$. Il polinomio risultante e' quindi facile da calcolare, sostituendo semplicemente nella formula gia' data:

$$\begin{aligned} e^x = \quad & 1 + \frac{1}{1!}x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \cdots + \frac{1}{n!}x^n + o(x^n) = \\ & 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!} + o(x^n) \end{aligned}$$

Esempio 6.10.2 *Calcolare, nel punto $x_0 = 0$, il polinomio di Taylor di ordine generico n della funzione:*

$$f(x) = \sin x$$

Questa volta abbiamo:

$$\begin{aligned} f^I(x) &= \cos x \\ f^{II}(x) &= -\sin x \\ f^{III}(x) &= -\cos x \\ f^{IV}(x) &= \sin x \end{aligned}$$

E da questo si deduce che ogni quattro derivate, il valore della derivata n-esima assume nuovamente quello della derivata (n-4)-esima. Dato che queste derivate vanno calcolate nel punto 0*, abbiamo la ripetizione di quattro valori:* 0, 1, 0, -1.

Il polinomio di Taylor di grado n-simo generico sara' dunque:

$$\sin x = 0 + \frac{1}{1!}x + 0 + \frac{-1}{3!}x^3 + 0 + \ldots$$

Si puo' notare facilmente che tutti i gradi pari sono nulli, e quindi non appaiono nel polinomio finale. Possiamo quindi prendere solo polinomi di grado dispari: nel caso di grado pari cambiera' solo l'errore rispetto al polinomio precendente, che si potra' prendere di un grado maggiore. Il risultato finale e' quindi:

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots + \frac{(-1)^k}{(2k+1)!}x^{2k+1} + o(x^{2k+1}), \quad k \in \mathbb{Z}$$

Esempio 6.10.3 *Calcolare, nel punto $x_0 = 0$, il polinomio di Taylor di ordine generico n della funzione:*

$$f(x) = \cos x$$

Anche in questo caso si arriva a conclusioni analoghe a quelle gia' formulate per il caso del seno. Passiamo direttamente al risultato finale (importante il fatto che, in questo caso, abbiamo solo gradi pari*.*

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots + \frac{(-1)^k}{(2k)!}x^{2k} + o(x^{2k})$$

Molto spesso i polinomi di Taylor vengono utilizzati nel calcolo di limiti, poiche' permettono di semplificare estremamente le espressioni.

Esempio 6.10.4 *Calcolare il seguente limite:*

$$\lim_{x\to 0} \frac{e^x - 1 - x - \frac{x^2}{2} + \sin x - x}{\cos x - 1}$$

Sviluppiamo le funzioni e^x, $\sin x$ e $\cos x$ utilizzando Taylor. Ma fino a che grado? Questo nessuno e' in grado di dircelo a priori: tutte le sostituzioni sono, a priori, accettabili, ma solo alcune combinazioni ci permettono di arrivare a risultati significativi, come vedremo. In questo caso e' semplice: infatti vediamo sottratto alla funzione e^x un polinomio di secondo grado che riconosciamo proprio come lo sviluppo di secondo grado di questa funzione, e simile cosa avviene per le funzioni seno e coseno in seguito. E' comunque buona norma partire con tentativi di basso grado, dove possibile di primo o secondo - raramente nella pratica si va oltre il terzo o addirittura quarto grado!!

$$\lim_{x\to 0} \frac{(1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + o(x^3)) - 1 - x - \frac{x^2}{2} + (x - \frac{x^3}{3!} + o(x^3)) - x}{(1 - \frac{x^2}{2} + o(x^2)) - 1}$$

Tra parentesi sono stati inseriti gli sviluppi di Taylor delle varie funzioni. Ora, semplificando, otteniamo:

$$\lim_{x\to 0} \frac{\frac{x^3}{3!} + o(x^3) - \frac{x^3}{3} - o(x^3)}{-\frac{x^2}{2} + o(x^2)}$$

A questo punto occorre dire un paio di parole sull'algebra degli o piccoli. C'e' infatti sempre, inanzitutto, da considerare il fatto che essi rappresentano una qualunque tra le possibili funzioni *che, divise per $(x-x_0)^n$, per $x \to x_0$, tendono a zero. Quindi, due o piccoli, pur essendo dello stesso grado, possono benissimo non essere lo stesso. In generale pero', si possono dimostrare queste asserzioni facilmente utilizzando la definizione stessa degli o piccoli (da notare che queste regole valgono solo per $x \to x_0 \in \mathbb{R}$, e non $\pm\infty$, ma d'altronde con le formule di Taylor non si possono maneggiare punti a piú o meno infinito):*

$$\begin{aligned}
&-o(x^n) = o(x^n)\\
&o(x^n) \pm o(x^n) = o(x^n)\\
&o(x^m) \pm o(x^n) = o(x^n); \quad m \geq n\\
&x^m + o(x^n) = o(x^n)\\
&x^m o(x^n) = o(x^{m+n})\\
&o(x^m)o(x^n) = o(x^{m+n})\\
&o(x^m + o(x^n)) = o(x^m); \quad m \leq n
\end{aligned}$$

Quindi nel nostro esempio, la differenza al numeratore tra i due o piccoli diventa un altro o piccolo:

$$\lim_{x\to 0} \frac{o(x^3)}{-\frac{1}{2}x^2 + o(x^2)}$$

A questo punto, dividiamo numeratore e denominatore per 0 (come sempre, poiché stiamo calcolando un limite, il punto $x = 0$ non ci interessa, e quindi possiamo dividere per 0 senza paura di dover considerare a parte il caso $x = 0$):

$$\lim_{x\to 0} \frac{o(x)}{-1/2 + o(1)}$$

Ora, dato che:

$$\lim_{x\to 0} o(x) = \lim_{x\to 0} \frac{o(x)}{x} x = 0 \cdot 0 = 0$$

Abbiamo che il nostro limite vale:

$$\frac{0}{-1/2} = 0$$

In questo esempio abbiamo percorso in modo pedante tutti i dovuti passaggi per arrivare alla soluzione: alla fine molti passaggi si salteranno perché automatici e con considerazioni sulla "velocità" con cui le funzioni tendono a 0.

Esempio 6.10.5 *Calcolare il polinomio di Taylor della funzione:*

$$f(x) = (1 + x)^\alpha$$

Dove $\alpha \in \mathbb{R}$.

Calcoliamo le prime derivate per tentare di ricavare l'andamento generale:

$$\begin{aligned}
&f'(x) = \alpha(1 + x)^{\alpha-1}\\
&f''(x) = \alpha(\alpha - 1)(1 + x)^{\alpha-2}\\
&\dots
\end{aligned}$$

Al che ricaviamo:

$$f(x) = 1 + \frac{\alpha}{1!}x + \frac{\alpha(\alpha-1)}{2!}x^2 + \frac{\alpha(\alpha-1)(\alpha-2)}{3!}x^3 + \cdots + \frac{\alpha(\alpha-1)(\alpha-2)\dots(\alpha-n+1)}{n!}x^n + o(x^n)$$

Ora, chiamiamo coefficiente binomiale generalizzato *la generalizzazione del normale coefficiente binomiale:*

$$\binom{p}{q} = \frac{\overbrace{p(p-1)\dots(p-q+1)}^{q \text{ volte}}}{q!}$$

Dove p può anche essere un numero reale. La nostra formula diventa ora:

$$(1+n)^\alpha = \sum_{i=0}^{n} \binom{\alpha}{i} x^i + o(x^n)$$

Esempio 6.10.6 *Utilizzando la formula precedente, calcolare al terz'ordine:*

$$\sqrt{1+x}$$

La soluzione è semplice: in questo caso abbiamo $\alpha = 1/2$, e quindi:

$$\begin{aligned}\sqrt{1+x} = \quad & 1 + \frac{1/2}{1!}x + \frac{1/2(1/2-1)}{2!}x^2 + \frac{1/2(1/2-1)(1/2-2)}{3!}x^3 + o(x^3) = \\ & 1 + \frac{x}{2} - \frac{x^2}{8} + \frac{x^3}{16} + o(x^3)\end{aligned}$$

A volte si può calcolare più facilmente il polinomio di Taylor della *derivata* della funzione cercata - in questo caso basterà cercare in seguito una funzione la cui derivata sia il polinomio trovato, e ciò per questo limitato caso è semplice (più avanti impareremo come la ricerca di *primitive* per qualunque funzione in realtà sia ben più complesso).

Esempio 6.10.7 *Calcolare il generico polinomio di Taylor della funzione:*

$$f(x) = ln(1+x)$$

Dato che

$$f'(x) = \frac{1}{1+x}$$

E di questo funzione sappiamo già che essa esprime il valore della serie geometrica (oppure si può arrivare allo stesso risultando considerando $(1+x)^{-1}$):

$$1 - x + x^2 - x^3 + \cdots + o(x^n)$$

E quindi, una funzione la cui derivata é quella appena data, risulta essere:

$$x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots + (-1)^n \frac{x^{n+1}}{n+1} + o(x^n)$$

Esempio 6.10.8 *Calcolare il polinomio di Taylor fino all'n-simo termine della funzione:*

$$f(x) = \arctan x$$

anche in questo caso, dato che sappiamo che:

$$f'(x) = \frac{1}{1+x^2} = 1 - x^2 + x^4 - \dots$$

Possiamo ricavare che:

$$f(x) = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots + (-1)^n \frac{x^{2n+1}}{2n+1}$$

Esempio 6.10.9 *Trovare per quali λ, $\mu \in \mathbb{R}$ si ha che:*

$$\lim_{x\to 0} \frac{\lambda\left(e^x - 1 - x - \frac{x^2}{2}\right) + \mu\left(\sin x - x\right)}{\arctan x - x} = 2$$

Anche in questo caso può rivelarsi utile lo sviluppo delle funzioni non polinomiali attraverso i polinomi di Taylor. Al denominatore possiamo sviluppare l'arcotangente solo per gradi dispari (basta guardare il suo sviluppo), quindi, superato il primo che viene eliminato da corrispondente "$-x$", si passa al terzo grado. Al denominatore abbiamo dunque un infinitesimo del terzo grado.

Al numeratore abbiamo la somma di due membri. Il primo contiene l'esponenziale a base e: possedendo essa tutti i gradi, basterà svilupparla fino al terzo grado. L'altro membro e' il seno, che possiede solo gradi dispari: anche in questo caso lo sviluppiamo fino al terzo grado.

$$\arctan x - x = x - \frac{x^3}{3} + o(x^3) - x = -\frac{x^3}{3} + o(x^3)$$

$$\lambda\left(e^x - 1 - x - \frac{x^2}{2}\right) = \lambda\left(1 + x + \frac{x^2}{2} + \frac{x^3}{6} + o(x) - 1 - x - \frac{x^2}{2}\right) = \lambda\left(\frac{x^3}{6} + o(x^3)\right)$$

$$\mu(\sin x - x) = \mu\left(x - \frac{x^3}{3} + o(x^3) - x\right) = \mu\left(-\frac{x^3}{3} + o(x^3)\right)$$

Adesso possiamo porre assieme tutti gli sviluppi che abbiamo trovato:

$$\lim_{x\to 0} \frac{\lambda\frac{x^3}{6} + o(x^3) - \mu\frac{x^3}{6} + o(x^3)}{-\frac{x^3}{6} + o(x^3)} =$$

$$\lim_{x\to 0} \frac{\frac{\lambda-\mu}{6}x^3 + o(x^3)}{-\frac{1}{3}x^3 + o(x^3)} = \lim_{x\to 0} \frac{\frac{\lambda-\mu}{6} + o(1)}{-\frac{1}{3} + o(1)} = -\frac{\lambda-\mu}{6}\cdot 3 = \frac{\mu-\lambda}{2}$$

Ora, dobbiamo porre questo limite uguale a 2:

$$\frac{\mu-\lambda}{2} = 2$$

Ovvero:

$$\mu = 4 + \lambda$$

Che e' la soluzione cercata.

Esempio 6.10.10 *Calcolare il seguente limite:*

$$\lim_{x\to 0} \frac{x^2\cos x + 1 - e^{x^2}}{x^2\sin^2 x}$$

Al numeratore abbiamo la somma di vari elementi. Esaminiamoli uno ad uno:

$$\cos x = 1 - \frac{x^2}{2} + o(x^2)$$

$$x^2\cos x = x^2 - \frac{x^4}{2} + o(x^4)$$

$$e^{(x^2)} = 1 + (x^2) + \frac{(x^2)^2}{2} + o\left((x^2)^2\right)$$

In questo ultimo caso, abbiamo sviluppato la funzione esponenziale considerando non la x come variabile, ma x^2. Il risultato è lo stesso, poichè, per $x \to 0$, anche $x^2 \to 0$, quindi è accettabile per l'uso con Taylor. Il risultato che quindi otteniamo è:

$$e^{x^2} = 1 + x^2 + \frac{x^4}{2} + o(x^4)$$

Quindi

$$1 - e^{x^2} = 1 - \left(1 + x^2 + \frac{x^4}{2} + o(x^4)\right) = -x^2 - \frac{x^4}{2} + o(x^4)$$

Mettendo assieme tutti i risultati ottenuti per il numeratore, otteniamo:

$$\begin{aligned} x^2\cos x + 1 - e^{x^2} &= x^2 - \frac{x^4}{2} + o(x^4) - x^2 - \frac{x^4}{2} + o(x^4) \\ &= -\frac{x^4}{2} + o(x^4) \end{aligned}$$

Al denominatore abbiamo il quadrato di un seno, moltiplicato per x^2. Tentiamo per cominciare di sviluppare il quadrato del seno al minor grado possibile, il primo:

$$\sin x = x + o(x)$$

Il suo quadrato sarà quindi:

$$\sin^2 x = (x + o(x))^2 = x^2 + 2xo(x) + o(x)^2$$

Utilizzando ora alcune delle regole aritmetiche sugli o piccoli:

$$x^2 + o(x^2) + o(x)o(x) = x^2 + o(x^2) + o(x^2) = x^2 + o(x^2)$$

Ora calcoliamo il rapporto:

$$\lim_{x\to 0} \frac{-x^4 + o(x^4)}{x^4 + o(x^4)} = -1$$

Capitolo 7

Integrali

7.1 Definizione di integrale

Presa una qualunque funzione $f(x)$ continua, si può ad essa associare un certo numero, indicato da simbolo:

$$\int_a^b f(x)\,\mathrm{d}x$$

La cui definizione è la seguente.

Dato un qualunque $n \in \mathbb{R} \geq 1$, decomponiamo l'intervallo $[a, b]$ in n intervallini, le cui ascisse chiameremo $x_0 = a$, x_1, x_2, $\ldots$, $x_n = b$. Inoltre, avremo sicuramente che

$$\forall j, 1 < j < n,\ x_j - x_{j-1} = \frac{b-a}{n}$$

Ovvero, la distanza tra due punti successivi è costante. Inoltre, al diminuire di n, questa lunghezza decresce. Ora, in ciascuno di questi intervallini $[x_{j-1}, x_j]$, fissiamo un ξ_j, stabilito in maniera arbitraria, ma tale che sia interno o su uno degli estremi dell'intervallo, ovvero:

$$x_{j-1} \leq \xi_j \leq x_j$$

Ed ora si considera la seguente somma:

$$f(\xi_1)(x_1 - x_0) + f(\xi_2)(x_2 - x_1) + \cdots + f(\xi_n)(x_n - x_{n-1}) = \sum_{j=1}^{n} f(\xi_j)(x_j - x_{j-1})$$

Questa somma, dato che abbiamo constatato che la lunghezza di ognuno di questi intervallini è costante, possiamo riscriverla come:

$$\frac{b-a}{n} \sum_{j=1}^{n} f(\xi_j) = \frac{b-a}{n}\ S_n\left(f; \xi = (\xi_0, \xi_1, \ldots, \xi_n)\right)$$

Dove il termine $S_n\left(f; \xi = (\xi_0, \xi_1, \ldots, \xi_n)\right)$ viene chiamato *somma parziale n-sima*, e rappresenta per l'appunto la corrispondente parte del membro sinistro dell'uguaglianza.

Ora, interrompiamo un attimo questa spiegazione, e passiamo ad una faccenda pratica. Definiamo così un insieme del piano:

$$A_f = \{(x, y) \in \mathbb{R} \mid a \leq x \leq b,\ 0 \leq y \leq f(x)\}$$

E cerchiamo di calcolarne l'area. Il modo più naturale che viene è quello di procedere per approssimazioni successive: ovvero, si divide l'intervallo $[a, b]$ in vari intervallini, e poi si approssima l'area di A_f in quella zona con un rettangolo. L'area totale approssimata sarà quindi proprio:

$$A_n = \frac{b-a}{n}\ S_n\left(f; \xi = (\xi_0, \xi_1, \ldots, \xi_n)\right)$$

Ecco quindi che abbiamo trovato un significato geometrico delle somme parziali.

Ora, l'area totale sarà il limite del valore di A_n per $n \to \infty$, ovvero:

$$A_f = \lim_{n\to\infty} \frac{b-a}{n}\ S_n\left(f; \xi = (\xi_0, \xi_1, \ldots, \xi_n)\right)$$

Questo limite, se esiste, viene chiamato *integrale di f tra a e b* e si scrive, per l'appunto:

$$\int_a^b f(x)\,\mathrm{d}x$$

Teorema 7.1.1 (Teorema fondamentale del calcolo integrale – 1^a parte) *Il limite*

$$\lim_{n\to\infty} \frac{b-a}{n} S_n\left(f; \xi = (\xi_0, \xi_1, \ldots, \xi_n)\right)$$

con $a, b \in \mathbb{R}$ ed f continua, converge indipendentemente dalla scelta della n-upla $\xi = (\xi_0, \xi_1, \ldots, \xi_n)$.

Questo teorema ci assicura quindi che l'integrale di una funzione continua su un intervallo esiste sempre – anche se il suo calcolo, come vedremo, può risultare assai complesso.

Ora possiamo dunque dare la definizione di area sottesa da una funzione:

$$\text{area}(A_f) \stackrel{def}{=} \int_a^b f(x)\,\mathrm{d}x$$

7.2 Proprietà degli integrali

Alcune proprietà elementari degli integrali, facilmente verificabili attraverso la definizione, sono le seguenti:

1. Linearità:
$$\int_a^b [\alpha f(x) + \beta g(x)]\,\mathrm{d}x = \alpha \int_a^b f(x)\,\mathrm{d}x + \beta \int_a^b g(x)\,\mathrm{d}x$$

2. Additività rispetto all'intervallo di integrazione:
$$\int_a^b f(x)\,\mathrm{d}x = \int_a^c f(x)\,\mathrm{d}x + \int_c^b f(x)\,\mathrm{d}x$$
Questa proprietà, per come abbiamo definito finora l'integrale, sarebbe corretta solo se $a \le c \le b$, ma in realtà grazie alla seguente:

3. Convenzione:
$$\int_a^b f(x)\,\mathrm{d}x = -\int_b^a f(x)\,\mathrm{d}x$$
Abbiamo che la proprietà precedente vale per qualunque a, b e c.

4. Monotonia:
$$\forall x \in [a,b] : f(x) \ge 0 \quad \Longrightarrow \quad \int_a^b f(x)\,\mathrm{d}x \ge 0$$
$$\forall x \in [a,b] : f(x) \ge g(x) \quad \Longrightarrow \quad \int_a^b f(x)\,\mathrm{d}x \ge \int_a^b g(x)\,\mathrm{d}x$$
E quindi, anche
$$\left|\int_a^b f(x)\,\mathrm{d}x\right| \le \int_a^b |f(x)|\,\mathrm{d}x$$

5. **Teorema 7.2.1 (Teorema della media)**
$$\exists c \in [a,b] \quad : \quad \frac{1}{b-a}\int_a^b f(x)\,\mathrm{d}x = f(c)$$

6. Funzioni pari
$$f(-x) = f(x) \Longrightarrow \int_{-a}^a f(x)\,\mathrm{d}x = 2\int_0^a f(x)\,\mathrm{d}x$$
Una funzione per cui vale che $f(-x) = f(x)$ si dice *funzione pari*.

7. Funzioni dispari
$$f(-x) = -f(x) \Longrightarrow \int_{-a}^a f(x)\,\mathrm{d}x = 0$$
Una funzione per cui vale che $f(-x) = -f(x)$ si dice *funzione dispari*. Il modo più intuitivo per dimostrare la validità di questi integrali è quello di pensare alla loro rappresentazione geometrica.

7.3 Teorema fondamentale del calcolo integrale

In generale, il calcolo del valore di un integrale per mezzo della definizione risulta piuttosto complicato. Infatti si tratta di trovare il limite di una funzione piuttosto complessa, una somma dal numero e valore di termini variabile. Tuttavia esiste un'altra via.

Inanzitutto bisogna introdurre il concetto di primitiva, già accennato quando si è parlato dei modi di risoluzione dei polinomi di Taylor. Si dice che una funzione F derivabile in $[a, b]$ è una *primitiva* di f in $[a, b]$ se $\forall x \in [a, b] : F'(x) = f(x)$.

É facile mostrare che se $F(x)$ è una primitiva di f, esistono infinite altre primitive e sono tutte e solo quelle rappresentabili nella forma $F(x) + k$, con k costante.

Non tutte le funzioni ammettono primitiva, tuttavia si può dimostrare che tutte le funzioni *continue* la ammettono.

Detto questo, possiamo enunciare:

Teorema 7.3.1 (Teorema fondamentale del calcolo integrale) *Se $f : [a, b] \to \mathbb{R}$ è continua, ed F è una sua primitiva su $[a, b]$; allora:*

$$\int_a^b f(x)\,\mathrm{d}x = F(b) - F(a)$$

Di solito la scrittura $F(b) - F(a)$ viene abbreviata con $[F(x)]_a^b$.

Questo teorema fornisce una risposta estremamente pratica (e allo stesso tempo, del tutto inaspettata!) al problema del calcolo di un integrale. La dimostrazione, comunque, è assai semplice.

Prendiamo n intervalli che dividono l'intervallo $[a, b]$: i punti che fanno ciò saranno dunque $x_0 = a$, x_1, ..., $x_n = b$. Ora:

$$F(b) - F(a) = F(x_n) - F(x_0) =$$
$$(F(x_n) - F(x_{n-1}) + (F(x_{n-1}) - F(x_{n-2}) + \cdots + (F(x_1) - F(x_0))$$

Per il teorema di Lagrange, presa una qualsiasi delle differenze $F(x_j) - F(x_{j-1})$, abbiamo che:

$$\exists \xi_j \in [x_j, x_{j-1}] : \frac{F(x_j) - F(x_{j-1})}{x_j - x_{j+1}} = F'(\xi_j)$$

Ma d'altronde, dato che F è una primitiva di f, abbiamo che $F'(\xi_j) = f(\xi_j)$, e quindi abbiamo:

$$\exists \xi_j \in [x_j, x_{j-1}] : F(x_j) - F(x_{j-1}) = f(\xi_j)(x_j - x_{j+1})$$

Quindi, l'uguaglianza iniziale diventa:

$$\exists (\xi_0, \xi_1, \ldots, \xi_n) : f(\xi_1)(x_1 - x_0) + f(\xi_2)(x_2 - x_1) + \cdots = S_n(f, \xi = (\xi_0, \xi_1, \ldots, \xi_n))$$

A questo punto, abbiamo ricavato proprio una delle somme parziali. Dato che il suo valore, qualunque sia la n, è lo stesso, ovvero $F(b) - F(a)$, per un'adeguata scelta di punti ξ (e questo ce l'assicura il teorema di Lagrange), il suo limite per $n \to \infty$ avrà sempre lo stesso valore, e quindi:

$$\int_a^b f(x)\,\mathrm{d}x = F(b) - F(a)$$

CVD.

7.4 Integrali indefiniti

Ora, quindi, il problema di calcolare un integrale (che, nella forma fino ad ora esaminata, prende il nome di *definito*) si è ridotto a quello di calcolare una primitiva di una funzione. L'operazione che indica il passaggio da una funzione alla sua primitiva sarà quindi una sorta di "inversa" della derivata, e prende il nome di *integrale indefinito*:

$$\int f(x)\,\mathrm{d}x$$

In realtà non si può trovare una vera e propria operazione inversa della derivata, in quanto esistono infinite funzioni che producono la stessa derivata, e sono tutte le funzioni che differiscono per una costante. Quindi, l'integrale indefinito non darà come risultato una singola funzione, bensì una famiglia di funzioni, tutte che differiscono per una costante.

Nel calcolo di un integrale definito il problema della presenza di questa costante scompare, infatti, se

$$\int f(x)\,\mathrm{d}x = F(x) + k$$

Allora

$$\int_a^b f(x)\,\mathrm{d}x = (F(b) + k) - (F(a) + k) = F(b) - F(a)$$

Quindi questa costante, qualunque sia, scompare.

Bisogna comunque sempre ricordare che, mentre il calcolo di una derivata, come abbiamo visto, è un'operazione meccanica e sempre possibile, a partire da funzioni esprimibili in termini di funzioni elementari, quello di un integrale non è sempre così semplice: anzi, spesso funzioni anche molto semplici, pur essendo continue ed ammettendo quindi integrale, non hanno integrali esprimibili elementarmente. Un classico esempio è quello di e^{-x^2}. Provare per credere: non esiste funzione che, derivata, dia questo risultato, eppure, essendo questa funzione continua, ammette integrale.

7.5 Integrali di potenze

Partiamo con l'esaminare uno tra i più semplici integrali che esistano: è facile mostrare che:

$$\int x^n \,\mathrm{d}x = \frac{x^{n+1}}{n+1} + k$$

Per convincersene, è sufficiente calcolare la derivata del secondo membro e ritrovare quindi il primo membro. Si noti come effettivamente sia stata aggiunta una costante k ad indicare una famiglia di funzioni.

In particolare si ha:

$$\int \mathrm{d}x = \int x^0 \,\mathrm{d}x = x$$

Quindi:

$$\int c \,\mathrm{d}x = c \int \mathrm{d}x = cx$$

E questo già ci consente, abbinando questa regola alla proprietà di linearità dell'integrale, di calcolare l'integrale di un qualunque polinomio.

Esempio 7.5.1 *Calcolare*

$$\int_1^2 (x^3 - 2x^2 + 5) \,\mathrm{d}x$$

Per mezzo di qualche semplice passaggio, abbiamo che:

$$\int_1^2 x^3 \,\mathrm{d}x - 2\int_1^2 x^2 \,\mathrm{d}x + 5\int_1^2 \mathrm{d}x = \left[\frac{x^4}{4}\right]_1^2 - 2\left[\frac{x^3}{3}\right]_1^2 + 5\,[x]_1^2 =$$

Quindi, risolvendo quelli che ormai sono solo calcoli aritmetici:

$$= \frac{16}{4} - \frac{1}{4} - 2\left(\frac{8}{3} + \frac{1}{3}\right) + 5\,(2-1) = \frac{11}{4}$$

7.6 Altri integrali immediati

Ovviamente, tutte le derivate elementari che abbiamo studiato possono semplicemente essere lette al contrario per ottenere i corrispondenti integrali elementari, con qualche semplificazione (vedi la tabella 7.6).

Tuttavia, mancano alcuni integrali che si potrebbero rivelare molto utili, come ad esempio $\ln x$ o $\tan x$: studiando i metodi di integrazione troveremo anche le risposte a queste domande.

A dire il vero, son ben pochi gli strumenti che abbiamo a nostra disposizione, e sono il metodo di *integrazione per parti* e quello di *integrazione per sostituzione*.

7.7 Integrazione per parti

Prendiamo due funzioni, $F(x)$ e $G(x)$, derivabili sull'intervallo I. Abbiamo allora, per la regola di derivazione di un prodotto di funzioni, che:

$$[F(x)G(x)]' = F'(x)G(x) + F(x)G'(x)$$

Ora, una primitiva del primo membro è, ovviamente, $F(x)G(x)$. Una primitiva del secondo membro la possiamo ricavare sommando le primitive dei suoi due addendi, ed otteniamo quindi:

$$F(x)G(x) = \int F'(x)G(x) \,\mathrm{d}x + \int F(x)G'(x) \,\mathrm{d}x$$

Ora, grazie a questa regola, conoscendo uno dei due integrali, posso calcolare anche l'altro per mezzo di una semplice differenza!

$$\int F'(x)G(x) \,\mathrm{d}x = F(x)G(x) - \int F(x)G'(x) \,\mathrm{d}x$$

È facile ricavare di conseguenza la analoga formula per gli integrali definiti, che è la seguente:

$$\int_a^b F'(x)G(x) \,\mathrm{d}x = [F(x)G(x)]_a^b - \int_a^b F(x)G'(x) \,\mathrm{d}x$$

$f(x)$	$\int f(x)\,\mathrm{d}x$
0	k
1	$x+k$
x^n	$\frac{x^{n+1}}{n+1}+k$
e^x	e^x+k
a^x	$\frac{a^x}{\ln a}+k$
$\cos x$	$\sin x+k$
$\sin x$	$-\cos x+k$
$\frac{1}{\cos^2 x}=1+\tan^2 x$	$\tan x+k$
$\sinh x$	$\cosh x$
$\cosh x$	$\sinh x$
$\frac{1}{\cosh^2 x}=1-\tanh^2 x$	$\tanh x$
$\frac{1}{x}$	$\ln\lvert x\rvert+k$
$\frac{1}{1+x^2}$	$\arctan x+k$
$\frac{1}{\sqrt{1-x^2}}$	$\arcsin x+k$
$-\frac{1}{\sqrt{1-x^2}}$	$\arccos x+k$
$\frac{1}{\sqrt{1+x^2}}$	$\operatorname{arcsinh} x$
$\frac{1}{\sqrt{x^2-1}}$	$\operatorname{arccosh} x$
$\frac{1}{1-x^2}$	$\operatorname{arctanh} x$

Tabella 7.1: Alcuni integrali dalle derivate elementari

Esempio 7.7.1 *Calcolare l'integrale:*

$$\int \ln x\,\mathrm{d}x$$

Per mezzo dell'integrazione per parti

Possiamo riscrivere l'integrale di partenza come:

$$\int (x)'\ln x\,\mathrm{d}x$$

E quindi, ora possamo usare la regola di integrazione per parti:

$$x\ln x-\int x(\ln x)'\,\mathrm{d}x = x\ln x-\int x\frac{1}{x}\,\mathrm{d}x =$$

E a questo punto, ormai il calcolo è finito:

$$= x\ln x-\int \mathrm{d}x = x\ln x-x+k = x(\ln x-1)+k$$

Esempio 7.7.2 *Calcolare*

$$\int x^n\ln x\,\mathrm{d}x$$

In questo caso l'integrazione per parti è ancora più evidente di prima:

$$\int\left(\frac{x^{n+1}}{n+1}\right)'\ln x\,\mathrm{d}x = \frac{x^{n+1}}{n+1}\ln x-\int\frac{x^n}{n+1}\,\mathrm{d}x$$

Ed anche in questo caso, ormai l'integrale è finito:

$$\frac{x^{n+1}}{n+1}\ln x-\frac{x^{n+1}}{(n+1)^2}+k = \frac{x^{n+1}}{n+1}\left(\ln x-\frac{1}{n+1}\right)+k$$

Esempio 7.7.3 *Calcolare*

$$\int \sqrt{1-x^2}\,\mathrm{d}x, \qquad x \in [-1;+1]$$

Si noti che questo non è tra gli integrali immediati da noi trovati.
Più avanti ripeteremo il calcolo di questo integrale col metodo di sostituzione, assai più semplice in questo caso.

$$\int \sqrt{1-x^2}\frac{\sqrt{1-x^2}}{\sqrt{1-x^2}}\,\mathrm{d}x = \int \frac{1-x^2}{\sqrt{1-x^2}}\,\mathrm{d}x = \int \frac{1}{\sqrt{1-x^2}}\,\mathrm{d}x - \int \frac{x^2}{\sqrt{1-x^2}}\,\mathrm{d}x$$

Il primo dei due integrali è presente tra gli integrali immediati, e vale $\arcsin x$*. Per calcolare il secondo invece facciamo questa osservazione:*

$$D\sqrt{1-x^2} = \frac{1}{2\sqrt{1-x^2}}(-2x) = -\frac{x}{\sqrt{1-x^2}}$$

E quindi, il secondo membro di quell'integrale risulta essere:

$$-\int \frac{x^2}{\sqrt{1-x^2}}\,\mathrm{d}x = \int D(\sqrt{1-x^2})x\,\mathrm{d}x$$

Ed ora possiamo di nuovo applicare il metodo di integrazione per parti a questo secondo integrale:

$$x\sqrt{1-x^2} - \int \sqrt{1-x^2}\,\mathrm{d}x$$

Dunque, ora, riprendendo l'integrale iniziale, siamo arrivati a questa uguaglianza:

$$\int \sqrt{1-x^2}\,\mathrm{d}x = \arcsin x + x\sqrt{1-x^2} - \int \sqrt{1-x^2}\,\mathrm{d}x$$

Ovvero, raccogliendo tutti gli integrali a sinistra:

$$2\int \sqrt{1-x^2} = \arcsin x + x\sqrt{1-x^2}$$

E quindi, dividendo per 2 *ed aggiungendo la costante:*

$$\int \sqrt{1-x^2} = \frac{\arcsin x + x\sqrt{1-x^2}}{2} + k$$

Questo integrale, seppur elaborato, mostra un altro modo per utilizzare il metodo per parti, ovvero riuscire a costruire un'uguaglianza in cui l'integrale di partenza appaia anche al secondo membro moltiplicato per una qualche costante diversa da 1 *(come* -1*, ad esempio, come è in questo caso), e quindi portarlo al primo membro e risolvere.*

Esempio 7.7.4 *Calcolare*

$$\int \sqrt{1+x^2}\,\mathrm{d}x$$

È facile mostrare, con ragionamenti analoghi a quelli già fatti, che questo integrale vale:

$$\frac{x\sqrt{1+x^2} + \operatorname{arcsinh} x}{2}$$

Esempio 7.7.5 *Calcolare*

$$\int \sin^2 x\,\mathrm{d}x$$

Anche in questo caso:

$$\int \sin x \cdot \sin x\,\mathrm{d}x = \int (-\cos x)' \sin x\,\mathrm{d}x = -\cos x \sin x + \int \cos x (\sin x)'\,\mathrm{d}x$$

Il secondo integrale diventa:

$$\int \cos^2 x\,\mathrm{d}x = \int (1-\sin^2 x)\,\mathrm{d}x = x - \int \sin^2 x\,\mathrm{d}x$$

Ora riscrivendo i due estremi dell'uguaglianza:

$$\begin{aligned}
\int \sin^2 x\,\mathrm{d}x &= -\cos x \sin x + x - \int \sin^2 x\,\mathrm{d}x \\
2\int \sin^2 x\,\mathrm{d}x &= x - \cos x \sin x \\
\int \sin^2 x\,\mathrm{d}x &= \frac{1}{2}(x - \cos x \sin x)
\end{aligned}$$

7.8 Integrazione per sostituzione

Sia F una primitiva di f in $[a;b]$. Sia ora $x=\varphi(t)$ su un intervallo $[\alpha;\beta]$ tale che

$$a=\varphi(\alpha) \qquad b=\varphi(\beta)$$

Per il teorema di derivazione di una funzione composta, abbiamo che:

$$\frac{\mathrm{d}F(\varphi(t))}{\mathrm{d}t}=\frac{\mathrm{d}F(\varphi(t))}{\mathrm{d}\varphi(t)}\cdot\frac{\mathrm{d}\varphi(t)}{\mathrm{d}t}=f(\varphi(t))\varphi'(t)$$

Dal che si deduce il seguente fatto: se $F(x)$ è una primitiva di $f(x)$, allora $\Phi(t)=G(\phi(t))$ è una primitiva di $f(\varphi(t))\cdot\varphi'(t)$. Da questo fatto, nasce la seguente regola di integrazione:

$$\int f(x)\,\mathrm{d}x=\int f(\varphi(t))\varphi'(t)\,\mathrm{d}t; \qquad x=\varphi(t)$$

Il che in pratica consiste semplicemente nel ricordare che $\mathrm{d}x=\mathrm{d}\varphi(t)=\varphi'(t)\,\mathrm{d}t$, come avevamo già notato trattando dei differenziali. Per l'integrale definito vale l'analoga regola:

$$\int_a^b f(x)\,\mathrm{d}x=\int_{\varphi^{-1}(a)}^{\varphi^{-1}(b)} f(\varphi(t))\varphi'(t)\,\mathrm{d}t$$

Si noti che in questo caso, al contrario dell'integrazione per parti, gli estremi dell'intervallo cambiano!

Esempio 7.8.1 *Calcolare*

$$\int\frac{\varphi'(t)}{\varphi(t)}\,\mathrm{d}t$$

Se poniamo

$$f(x)=\frac{1}{x}$$

Abbiamo:

$$\frac{\varphi'(t)}{\varphi(t)}=\varphi'(t)\frac{1}{\varphi(t)}=f(\varphi(t))\varphi'(t)$$

E quindi

$$\int f(\varphi(t))\varphi'(t)\,\mathrm{d}t=\int f(\varphi(t))\,\mathrm{d}\varphi(t)$$

Ma dato che

$$\int f(x)\,\mathrm{d}x=\int\frac{1}{x}\,\mathrm{d}x=\ln|x|+k$$

Allora

$$\int f(\varphi(t))\,\mathrm{d}\varphi(t)=\ln|\varphi(t)|+k$$

Esempio 7.8.2 *Calcolare*

$$\int\tan x\,\mathrm{d}x$$

Se trasformiamo l'integrale come segue:

$$\int\frac{\sin x}{\cos x}\,\mathrm{d}x=-\int\frac{(\cos x)'}{\cos x}\,\mathrm{d}x=-\ln|\cos x|+k$$

Utilizzando la regola ricavata dall'esercizio precedente.

Esempio 7.8.3 *Calcolare*

$$\int\arctan x\,\mathrm{d}x$$

Anche in questo caso, con qualche trasformazione, ci riconduciamo al caso iniziale; prima applichiamo una integrazione per parti:

$$\int(x)'\arctan x\,\mathrm{d}x=x\arctan x-\int\frac{x}{x^2+1}\,\mathrm{d}x=x\arctan x-\frac{1}{2}\int\frac{2x}{x^2+1}\,\mathrm{d}x$$

E a questo punto possiamo applicare l'integrazione per sostituzione:

$$x\arctan x - \frac{1}{2}\int \frac{1}{x^2+1}\,\mathrm{d}(x^2+1) = x\arctan x - \frac{1}{2}\ln(1+x^2) + k$$

Si può notare il fatto curioso che l'integrale dell'arcotangente abbia nella soluzione un logaritmo. In realtà i collegamenti tra le funzioni trigonometriche ed esponenziali (e tra le loro rispettive inverse) sono vari, e spiegabili se le si esaminano nel campo complesso, ma questo esula dall'argomento corrente.

Esempio 7.8.4 *Calcolare*

$$\int \arcsin x\,\mathrm{d}x$$

Utilizzando un procedimento analogo al precedente:

$$\int (x)'\arcsin x\,\mathrm{d}x = x\arcsin x - \int \frac{x}{\sqrt{1-x^2}}\,\mathrm{d}x$$

A questo punto notiamo che:

$$(\sqrt{1-x^2})' = \frac{1}{x\sqrt{1-x^2}}\cdot(-2x) = -\frac{x}{\sqrt{1-x^2}}$$

E quindi, l'integrale di partenza vale:

$$x\arcsin x + \sqrt{1-x^2} + k$$

La formula trovata in realtà si può anche utilizzare in senso inverso, ovvero riscrivere una certa funzione di x come variabile indipendente, come mostrano i seguenti esempi.

Esempio 7.8.5 *Calcolare*

$$\int \frac{1}{e^x + e^{-x}}\,\mathrm{d}x$$

Se si pone:

$$\begin{aligned} e^x &= t \\ x &= \ln t \\ \mathrm{d}x &= \frac{\mathrm{d}t}{t} \end{aligned}$$

E quindi:

$$\int \frac{1}{t+\frac{1}{t}}\frac{\mathrm{d}t}{t} = \int \frac{1}{1+t^2}\,\mathrm{d}t = \arctan t + k = \arctan e^x + k$$

7.9 Integrali impropri

Spesso si può presentare il caso in cui la funzione che si tenta di integrare non sia continua. In questo caso, quando la funzione è comunque integrabile, e quale è il suo integrale?

Il caso piu' semplice è quello in cui la funzione presenti un numero *finito* di salti. In tal caso basta utilizzare la proprietà di additività rispetto all'intervallo di integrazione, e dividere l'integrale nelle varie parti in cui la funzione rimane continua. In tal modo si ha la somma di più integrali di funzioni continue.

Un caso più complesso è quello in cui la funzione in un qualche punto diverga a $\pm\infty$.

Sia $f : (a;b] \to \mathbb{R}$ continua, e $\lim_{x\to\infty} f(x) = \pm\infty$. In questo caso ovviamente non si può usare la definizione data finora, dato che questo integrale non è calcolato su una funzione continua. Tuttavia, se prendiamo l'intervallo $[a+\varepsilon; b]$, $\varepsilon > 0$, qui la f è continua, e qui l'integrale siamo certi che si possa calcolare:

$$\int_{a+\varepsilon}^{b} f(x)\,\mathrm{d}x$$

A questo punto, poniamo *per definizione*:

$$\int_a^b f(x)\,\mathrm{d}x \overset{def}{=} \lim_{\varepsilon\to 0^+}\int_{a+\varepsilon}^{b} f(x)\,\mathrm{d}x$$

A seconda del risultato del limite proposto, anche l'integrale ha valori diversi:

- Se il limite è finito, si dice che l'integrale esiste, o che è *convergente*.
- Se vale $\pm\infty$, si dice che la funzione non è integrabile, o che l'integrale è *divergente*.
- Se non esiste, non esiste integrale

Analogamente, per un intervallo $[a;b)$ dove f è continua e $\lim_{x\to b} f(x) = \pm\infty$, si pone per definizione:

$$\int_a^b f(x)\,\mathrm{d}x \stackrel{def}{=} \lim_{\varepsilon\to 0^+} \int_a^{b-\varepsilon} f(x)\,\mathrm{d}x$$

Infine, se ad entrambi gli estremi la funzione diverge, si prende un qualunque $T \in (a;b)$ e si pone:

$$\int_a^b f(x)\,\mathrm{d}x = \int_a^T f(x)\,\mathrm{d}x + \int_T^b f(x)\,\mathrm{d}x$$

Questo implica che *entrambi* gli integrali esistano per permettere l'esistenza dell'integrale di partenza.

Ovviamente, così come si è trovato il modo per calcolare un integrale in un punto in cui la funzione divergeva ad infinito, si può anche trovare il modo di calcolare l'integrale se l'intervallo preso in considerazione è non finito. Ad esempio, presa $f : [a;+\infty) \to \mathbb{R}$ ivi continua, si pone:

$$\int_a^{+\infty} f(x)\,\mathrm{d}x = \lim_{\varepsilon\to+\infty} \int_a^{\varepsilon} f(x)\,\mathrm{d}x$$

Le definizioni analoghe sono simmetriche a quelle già date nell'altro caso.

7.9.1 Criteri di integrabilità

Spesso occorre semplicemente sapere se un certo integrale esiste ed è finito, ma senza la necessità di calcolarlo (cosa, spesso, impossibile). Per vedere ciò, esistono i cosiddetti *criteri di integrabilità*. Ma prima, calcoliamo l'integrale improprio di alcune funzioni importanti.

Esempio 7.9.1 *Calcolare*

$$\int_a^b \frac{1}{(x-a)^\alpha}\,\mathrm{d}x, \qquad \alpha > 0$$

Inanzitutto osserviamo che in a è presente una discontinuità, quindi dobbiamo utilizzare la definizione di integrale improprio:

$$\lim_{\varepsilon\to 0^+} \int_{a+\varepsilon}^b \frac{1}{(x-a)^\alpha}\,\mathrm{d}x$$

Calcoliamo l'integrale inanzitutto.

$$\int (x-a)^\alpha\,\mathrm{d}x = \begin{cases} \alpha = 1: & \ln|x-a| \\ \alpha \neq 1: & \dfrac{(x-a)^{1-\alpha}}{1-\alpha} \end{cases}$$

Ora calcoliamo il limite nei due casi:

$$\lim_{\varepsilon\to 0^+} \left[\ln|x-a|\right]_{a+\varepsilon}^b = \lim_{\varepsilon\to 0^+} (\ln|b-a| - \ln\varepsilon) = \infty$$

Quindi, per $\alpha = 1$, la funzione non è integrabile. Nel secondo caso abbiamo invece:

$$\lim_{\varepsilon\to 0^+} \left[\frac{(x-a)^{1-\alpha}}{1-\alpha}\right]_{a+\varepsilon}^b = \lim_{\varepsilon\to 0^+} \frac{(b-a)^{1-\alpha}}{1-\alpha} - \frac{1}{1-\alpha}\varepsilon^{1-\alpha}$$

In pratica, tutto si è ridotto a calcolare il limite

$$\lim_{\varepsilon\to 0^+} \varepsilon^{1-\alpha}$$

Il quale è finito se $1-\alpha > 0$, ovvero se $0 < \alpha < 1$, mentre è infinito se $1-\alpha < 0$, ovvero se $\alpha > 1$. Riassumento il tutto, abbiamo dimostrato che l'integrale dato esiste finito per $0 \leq \alpha < 1$, mentre diverge per $\alpha > 1$.

A questo punto possiamo enunciare il seguente:

Criterio 7.9.1 (Criterio di integrabilità al finito) *Sia $f : (a;b] \to \mathbb{R}$, continua su questo intervallo, e per cui $\lim_{x\to a} f(x) = \pm\infty$. Allora, se $f : (a;b] > 0$, $g : (a;b] > 0$, $\exists \alpha < 1 : f \sim \dfrac{1}{(x-a)^\alpha}$, f è integrabile.*

In generale, se non si riesce in modo diretto, si può sempre utilizzare un'approssimazione per mezzo dei polinomi di Taylor della funzione per dimostrare la sua asintoticità con $1/(x-a)^\alpha$. In generale, la condizione appena data equivale a:

$$\exists \alpha < 1 \ : \ \lim_{x\to a}(x-a)^\alpha |f(x)| = k \in \mathbb{R}$$

Analogamente, esiste anche il criterio opposto:

Criterio 7.9.2 (Criterio di non integrabilità al finito) *Se*

$$\exists \alpha \geq 1 \ : \ \lim_{x\to a}(x-a)^\alpha f(x) = k \in \mathbb{R} \neq 0$$

Allora la funzione data non *è integrabile.*

Se si osserva con attenzione, vi sono un paio di asimmetrie: queste sono dovute al fatto che è presente, oltre alla possibilità che la funzione sia o meno integrabile, anche una terza, ovvero che l'integrale non esista.

Entrambi questi criteri presentano analoghe forme per il punto di infinito è b e non a.

Esempio 7.9.2 *Determinare se*

$$\int_0^1 \frac{1}{\sqrt{1-x^3}}\,\mathrm{d}x$$

esiste finito.

Osserviamo che il punto di discontinuità è 1*. Quindi, dobbiamo vedere se esiste un* $\alpha < 1$ *tale per cui:*

$$\lim (1-x)^\alpha \left|\frac{1}{\sqrt{1-x^3}}\right| = \lim \frac{(1-x)^\alpha}{\sqrt{1-x^3}}$$

esista finito. D'altronde, la funzione sotto limite la possiamo riscrivere come:

$$\frac{(1-x)^\alpha}{\sqrt{(1-x)(1+x+x^2)}} = \frac{(1-x)^{\alpha-\frac{1}{2}}}{\sqrt{1+x+x^2}}$$

Ora, se si passa al limite, il denominatore tende a $\sqrt{3}$*, quindi, possiamo ad esempio prendere* $\alpha = \frac{1}{2}$ *per avere limite finito. Dato che questo valore di* α *è minore di* 1*, abbiamo dimostrato che l'integrale esiste finito.*

Esempio 7.9.3 *Determinare se*

$$\int_0^{\pi/2} \frac{\mathrm{d}x}{1-\cos x}$$

esiste finito.

Dato che la funzione integranda presenta una discontinuità nel punto 0*, utilizziamo uno dei criteri a nostra disposizione per verificare se è integrabile o meno. Da notare che nell'intervallo* $[0;\pi/2]$ *la funzione è sempre positiva, quindi non occorrerà segnalare il valore assoluto nei calcoli. Quindi il nostro scopo è quello di determinare un* α *per cui*

$$\lim_{x\to 0} \frac{x^\alpha}{1-\cos x}$$

Sia finito. Per trovarlo basta scomporre il limite come segue e ricordare un semplice limite notevole:

$$\lim_{x\to 0} \frac{x^\alpha}{x^2}\frac{x^2}{1-\cos x} = \frac{1}{2}\lim_{x\to 0}\frac{x^\alpha}{x^2}$$

Che tende ad un limite finito per $\alpha \geq 2$*. Per la precisione, tende ad un limite finito diverso da zero per* $\alpha = 2$*, e quindi, per il criterio di non integrabilità, l'integrale dato non è integrabile.*

Da notare anche che, a parte questo metodo, si poteva molto semplicemente utilizzare la definizione di integrale impoprio per verificarne l'integrabilità o meno. Infatti:

$$\int_0^{\pi/2} \frac{\mathrm{d}x}{1-\cos x} = \lim_{\varepsilon\to 0^+}\int_\varepsilon^{\pi/2}\frac{\mathrm{d}x}{1-\cos x}$$

Ed ora calcoliamo l'integrale indefinito dato, ricordando che:

$$\cos x = \cos\left(2\frac{x}{2}\right) = 1 - 2\sin^2\frac{x}{2}$$

E quindi

$$\int \frac{\mathrm{d}x}{1-\cos x} = \int \frac{\mathrm{d}x}{1-\left(1-2\sin^2\frac{x}{2}\right)} = \int \frac{2\,\mathrm{d}\frac{x}{2}}{2\sin^2\frac{x}{2}} = -\cot\frac{x}{2} + k$$

Il limite varrà dunque:

$$\lim_{\varepsilon\to 0^+}\left[-\cot\frac{x}{2}\right]_\varepsilon^{\pi/2} = \lim_{\varepsilon\to 0^+}\left(\cot\frac{\varepsilon}{2} - 1\right) = +\infty$$

Ecco ritrovato il risultato già ricavato, ovvero che l'integrale non esiste finito.

La base dei criteri appena enunciati è la seguente:

Criterio 7.9.3 (Criterio di integrabilità: confronto asintotico al finito) *Sia* $f, g : (a; b] \to \mathbb{R}$, $\forall x \in (a; b] : f(x) > 0, g(x) > 0$, $f \sim g$ *per* $x \to a^+$, *allora* f *è integrabile se e solo se* g *è integrabile su* $(a; b]$.

Questo ovviamente non assicura che gli integrali di f e g abbiano lo stesso valore, anzi, molto spesso non è così!

Come si può notare il criterio è analogo a quello di convergenza delle serie – dobbiamo d'altronde ricordare l'affinità tra i due concetti: l'integrale è il limite di una particolare serie. In effetti, esistono anche altri criteri analoghi a quelli delle serie validi per gli integrali:

Criterio 7.9.4 (Criterio di integrabilità: confronto al finito) *Sia* $f, g : (a; b] \to \mathbb{R}$, $\forall x \in (a; b] : 0 \leq f(x) \leq g(x)$, *allora* g *integrabile* $\Rightarrow$ f *integrabile, e* f *non integrabile* $\Rightarrow$ g *non integrabile.*

Questo criterio si dimostra facilmente per mezzo della proprietà di monotonia degli integrali.

Criterio 7.9.5 (Criterio di integrabilità assoluta al finito)

$$\int_a^b |f(x)| \, \mathrm{d}x \text{ convergente} \Longrightarrow \int_a^b f(x) \, \mathrm{d}x \text{ convergente}$$

Se di una funzione è integrabile il suo valore assoluto, allora si dice che è *assolutamente integrabile.*

Per integrali su intervalli non finiti valgono simili criteri, come quello di confronto, confronto asintotico, integrabilità assoluta. I casi particolari dei criteri di confronto asintotici presentati ovviamente devono cambiare.

Inanzitutto si può dimostrare attraverso la definizione che

$$\frac{1}{x^{\alpha>1}}$$

è integrabile, sull'intervallo $[a; +\infty]$. Da qui, otteniamo:

Criterio 7.9.6 (Criterio di integrabilità) *Se*

$$\exists \alpha > 1 : \lim_{x \to +\infty} x^\alpha |f(x)| = k \in \mathbb{R}$$

Allora l'integrale

$$\int_a^{+\infty} f(x) \, \mathrm{d}x$$

Esiste finito.

Ed abbiamo anche

Criterio 7.9.7 (Criterio di non integrabilità) *Se*

$$\exists 0 \leq \alpha < 1 : \lim_{x \to +\infty} x^\alpha f(x) = k \in \mathbb{R} \neq 0$$

Allora l'integrale

$$\int_a^{+\infty} f(x) \, \mathrm{d}x$$

Non *esiste finito.*

Esempio 7.9.4 *Dimostrare che esiste finito*

$$\int_{-\infty}^{+\infty} e^{-x^2} \, \mathrm{d}x$$

(detto integrale di Poisson*).*

Quindi dobbiamo dimostrare che esistono finiti, ad esempio, gli integrali

$$\int_0^{+\infty} e^{-x^2} \, \mathrm{d}x$$

e

$$\int_{-\infty}^{0} e^{-x^2} \, \mathrm{d}x$$

Ovviamente, se l'uno esiste finito, anche l'altro lo è. Quindi dobbiamo trovare un α *per cui*

$$\lim_{x \to +\infty} x^\alpha e^{-x^2}$$

Esiste finito. Ma questo limite tende a zero proprio per ogni $\alpha > 1$, *infatti:*

$$\lim_{x \to +\infty} x^\alpha e^{-x^2} = \lim_{x \to +\infty} \frac{x^\alpha}{e^{x^2}}$$

E questo è un rapporto tra due infiniti, ed al denominatore ne è presente uno di ordine maggiore.

Per la precisione si può dimostrare che questo integrale vale $\sqrt{\pi}$.

Non sempre d'altronde i due criteri finora presentati sono sufficienti a dimostrare l'integrabilità o meno di alcune funzioni.

Esempio 7.9.5 *Determinare, se possibile, l'integrabilità di*

$$\int_1^{+\infty} \frac{\sin x}{x}\,\mathrm{d}x$$

Utilizzando il criterio di integrabilità, dovremmo cercare il limite di

$$\lim_{x\to\infty} x^\alpha \frac{|\sin x|}{x} = \lim_{x\to\infty} x^{\alpha-1}|\sin x|$$

Non si può usare il criterio di integrabilità, poichè per ogni $\alpha > 1$ questo limite è non finito. Ma non si può usare neanche il criterio di non integrabilità, poichè per $0 < \alpha < 1$ il limite tende a 0. Quindi, dato che la funzione integranda non ha integrale esprimibile sotto forma di funzioni elementari, non abbiamo strumenti, apparentemente, per determinare la convergenza o meno dell'integrale.

Tuttavia una via esiste comunque. Prima operiamo alcune trasformazioni sull'integrale:

$$\int_1^{+\infty} \frac{\sin x}{x}\,\mathrm{d}x = \lim_{T\to+\infty} \int_1^T \frac{\sin x}{x}\,\mathrm{d}x$$

E questo punto, come sempre, operiamo su questo integrale.

$$\int_1^T \frac{(-\cos x)'}{x}\,\mathrm{d}x = \Big[-\frac{\cos x}{x}\Big]_1^T - \int_1^T \frac{\cos x}{x^2}\,\mathrm{d}x = -\frac{\cos T}{T} + \cos 1 - \int_1^T \frac{\cos x}{x^2}\,\mathrm{d}x$$

Se calcoliamo il limite di questo integrale, dato che $\lim \cos T/T = 0$, allora avremo risultato finito solo se sarà finito il limite dell'integrale, e quindi abbiamo che

$$\int_1^{+\infty} \frac{\sin x}{x}\,\mathrm{d}x \text{ converge} \iff \int_1^{+\infty} \frac{\cos x}{x^2}\,\mathrm{d}x \text{ converge}$$

Apparentemente, sembra la stessa situazione di prima, anzi, più complessa, ma in realtà non è così. Infatti, proviamo ad utilizzare i criteri a nostra disposizione su questo nuovo integrale:

$$\frac{|\cos x|}{x^2} x^\alpha = |\cos x| x^{\alpha-2}$$

Questo limite è finito se l'esponente della x è minore di 0. Per il criterio di integrabilità, esiste effettivamente almeno un valore di α maggiore di 1 (tutti gli $1 < \alpha < 2$ soddisfano questa condizione) per cui il limite è finito (0), e quindi questa funzione è integrabile. Per la relazione trovata prima, anche la funzione di partenza è integrabile.

7.10 Integrali di rapporti di polinomi

Un problema che spesso si può presentare è quello di integrare il rapporto di due polinomi, ovvero di calcolare un integrale nella seguente forma:

$$\int \frac{P_n(x)}{Q_m(x)}\,\mathrm{d}x$$

Dove P_n e Q_m sono polinomi di grado n ed m, rispettivamente.

Anzitutto: se il grado del numeratore è maggiore di quello del denominatore, ovvero $n > m$, si esegue la divisione tra i due polinomi, ottenendo:

$$\frac{P_n(x)}{Q_m(x)} = D_{n-m}(x) + \frac{R_k(x)}{Q_m(x)}$$

Dove si è sicuri che $k < m$. Dato che sappiamo già calcolare l'integrale di un qualunque polinomio come D, il problema si è ricondotto a quello di calcolare l'integrale di un rapporto di polinomi, *in cui il numeratore ha grado minore del denominatore.*

Ora cominciamo a considerare alcuni casi particolari.

7.10.1 Caso $P_0(x)/Q_1(x)$

Ovvero quando al numeratore è presente una costante ed al denominatore un polinomio di primo grado in x. È il caso più semplice, in quanto l'integrale è calcolabile con una semplice sostituzione, ed il risultato è un logaritmo.

Esempio 7.10.1 *Calcolare*

$$\int \frac{2}{3x-4}\,\mathrm{d}x$$

Ponendo

$$t = 3x - 4$$

Si ottiene

$$2\int \frac{1}{t}\frac{1}{3}\,\mathrm{d}t = \frac{2}{3}\ln|t| + k = \frac{2}{3}\ln|3x-4| + k$$

7.10.2 Caso $P_0(x)/Q_2(x)$

In questo caso, a seconda del valore del determinante Δ del denumeratore, bisogna distinguere in tre diversi casi:

- $\Delta > 0$

  In questo caso, il denominatore ax^2+bx+c si potrà scrivere come prodotto di due polinomi di primo grado $((x-x_1)(x-x_2))$, e quindi l'intera frazione si potrà scrivere come somma di due frazioni aventi al denominatore polinomi di primo grado in x.

  Esempio 7.10.2 *Calcolare*

  $$\int \frac{3}{x^2-3x+2}$$

  Il denominatore ha discriminante $\Delta = 3\cdot 3 - 4\cdot 2 = 1 > 0$*, quindi troviamo le soluzioni dell'equazione* $Q(x) = 0$*, che sono* $x = 1$ *ed* $x = 2$*, quindi* $x^2 - 3x + 2 = (x-2)(x-1)$*. Ora possiamo scrivere che*

  $$\frac{1}{x^2-3x+2} \equiv \frac{A}{x-1} + \frac{B}{x-2}$$

  (la costante 3 *della frazione originaria è stata eliminata in quanto può essere subito portata fuori dal segno di integrale).*

  Ora è sufficiente svolgere i calcoli per trovare A *e* B*:*

  $$\frac{1}{x^2-3x+2} \equiv \frac{A(x-1)+B(x-2)}{(x-2)(x-1)}$$

  $$1 \equiv (A+B)x + (A-2B)$$

  $$\begin{cases} A+B=0 \\ A-2B=1 \end{cases}$$

  $$\begin{cases} A=-B \\ -3B=1 \end{cases}$$

  $$\begin{cases} A=\dfrac{1}{3} \\ B=-\dfrac{1}{3} \end{cases}$$

  Ecco che abbiamo trovato questi valori. L'integrale di partenza risulta quindi essere

  $$\int \frac{3}{x^2-3x+2} = 3\left(\int \frac{1}{3}\frac{1}{x-2}\,\mathrm{d}x - \int \frac{1}{3}\frac{1}{x-1}\,\mathrm{d}x\right)$$

  Ed ora sono solo calcoli di routine.

  $$\ln|x-2| - \ln|x-1| = \ln\left|\frac{x-2}{x-1}\right|$$

- $\Delta = 0$ In questo caso si può riscrivere il denominatore come quadrato perfetto, e con una semplice sostituzione, arrivare al risultato.

Esempio 7.10.3 *Calcolare*

$$\int \frac{1}{4x^2 - 12x + 9}\,\mathrm{d}x$$

Il denominatore ha determinante nullo, ed in effetti si può riscrivere nella forma $(2x-3)^2$*. Se ora pongo* $t = 2x - 3$*, ottengo:*

$$\int \frac{1}{t^2}\frac{\mathrm{d}t}{2} = \frac{1}{2}\int t^{-2}\,\mathrm{d}t = -\frac{1}{2t} + k = -\frac{1}{2(2x-3)} + k$$

- $\Delta < 0$ In questo caso si usa la cosiddetta *tecnica del completamento del quadrato*, ovvero si trasforma il polinomio nella somma di un quadrato ed un numero. A quel punto, per mezzo di una sostituzione, ci si riconduce ad una arcotangente.

Esempio 7.10.4 *Calcolare*

$$\int \frac{1}{x^2 - 2x + 3}$$

Il determinante del denominatore è minore di zero. Per usare la tecnica del quadrato, devo riuscire a costruire un quadrato perfetto a partire dal polinomio di secondo grado. Questo quadrato perfetto avrà sicuramente come coefficiente della x *il valore* 1*, poichè* $(1 \cdot x)^2 = x^2$*, che è la forma con cui appare il termine di secondo grado. Poi, il doppio prodotto dei coefficienti deve valere* -2*, quindi, se chiamiamo* a *il coeffiente da noi cercato,* $2 \cdot 1 \cdot a = -2$*, ovvero* $a = -1$*. Quindi il quadrato da noi trovato vale* $(x-1)^2$*. Ma questo da come risultato* $x^2 - 2x + 1$*, mentre noi abbiamo* $x^2 - 2x + 3$*, quindi dobbiamo scrivere:*

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

A questo punto il denominatore si può scrivere come somma di due quadrati, ovvero i quadrati di $x - 1$ *e di* $\sqrt{2}$*. Quindi, se applichiamo il la sostituzione* $t = x - 1$ *abbiamo*

$$\int \frac{1}{t^2 + (\sqrt{2})^2}\,\mathrm{d}x$$

Quindi il nostro problema si è ora ricondotto a quello di calcolare l'integrale di un integrale del tipo:

$$\int \frac{1}{x^2 + a^2}$$

Utilizzando la sostituzione $x = at$*, da cui* $\mathrm{d}x = a\,\mathrm{d}t$*, abbiamo:*

$$\int \frac{a\,\mathrm{d}t}{a^2t^2 + a^2} = \frac{1}{a}\int \frac{1}{1+t^2}\,\mathrm{d}t = \frac{1}{a} + \arctan\frac{x}{a} + k$$

Quindi, tenendo a mente questo risultato, il nostro integrale ora diventa:

$$\int \frac{1}{t^2 + (\sqrt{2})^2}\,\mathrm{d}x = \frac{1}{\sqrt{2}}\arctan\frac{t}{\sqrt{2}} + k = \frac{1}{\sqrt{2}}\arctan\frac{x-1}{\sqrt{2}} + k$$

Ed ecco trovato il nostro risultato.

7.10.3 Caso $P_1(x)/Q_2(x)$

In questo caso occorre trasformare il numeratore nella derivata del denominatore e poi utilizzare un cambio di variabile.

Esempio 7.10.5 *Calcolare*

$$\int \frac{2x+1}{3x^2 - x + 1}\,\mathrm{d}x$$

La derivata del denominatore è $6x - 1$*. Partiamo col mettere a posto il coefficiente del primo grado del numeratore, che ora è* 2 *– occorre quindi moltiplicare per* 3 *all'interno dell'integrale.*

$$\frac{1}{3}\int \frac{3(2x+1)}{3x^2 - x + 1}\,\mathrm{d}x = \frac{1}{3}\int \frac{6x+3}{3x^2 - x + 1}\,\mathrm{d}x$$

Ora invece bisogna sistemare il termine costante, e si può fare semplicemente aggiungendo e togliendo -1*:*

$$\frac{1}{3}\int \frac{6x+3-1+1}{3x^2 - x + 1}\,\mathrm{d}x = \frac{1}{3}\left(\int \frac{6x-1}{3x^2 - x + 1}\,\mathrm{d}x + \int \frac{4}{3x^2 - x + 1}\,\mathrm{d}x\right)$$

Ora, il primo di questi integrali lo calcoliamo con la semplice sostituzione $t = 3x^2 - x + 1$*, e quindi* $\mathrm{d}t = (6x-1)\,\mathrm{d}x$*:*

$$\int \frac{(6x-1)\,\mathrm{d}x}{3x^2-x+1} = \int \frac{\mathrm{d}t}{t} = \ln|t| + k = \ln(3x^2-x+1) + k$$

Il secondo integrale è invece di uno dei tipi visti precedentemente: in questo caso $\Delta < 0$*, e quindi compiamo un completamento del quadrato. Il coefficiente della* x *sarà* $\sqrt{3}$*, poichè* $(\sqrt{3}x)^2 = 3x^2$*. poi,* $-x = 2 \cdot \sqrt{3} \cdot a \cdot x$*, quindi* $a = -1/2\sqrt{3}$*. Il quadrato così ottenuto si svilupperebbe però in* $x^2 - x + 1/12$*, mentre noi abbiamo* $3x^2 - x + 1$*, e quindi:*

$$\int \frac{4}{3x^2-x+1}\,\mathrm{d}x = 4\int \frac{1}{\left(\sqrt{3}x - \frac{1}{2\sqrt{3}}\right)^2 + \frac{11}{12}}$$

A questo punto attraverso una sostituzione e la regola trovata precedentemente, si ottiene per questo secondo integrale il risultato:

$$\frac{8}{\sqrt{11}} \arctan\left[\frac{\sqrt{11}}{2}\left(x - \frac{1}{6}\right)\right]$$

Quindi, a questo punto basta sommare i due integrali trovati e si ha il risultato.

7.11 Volumi di solidi di rotazione

Una breve nota riguardo ai volumi dei cosiddetti *solidi di rotazione*. Vengono chiamati così quei solidi che si possono costruire partendo da una curva e facendola ruotare intorno ad un asse. Ad esempio il cilindro, il cono, la sfera, sono tutti solidi di rotazione.

Prendiamo un solido costruito facendo ruotare intorno all'asse x il tratto di grafico di una funzione continua $f(x)$ presa nell'intervallo $[a;b]$. Intuitivamente (per una dimostrazione più precisa occorrerà prima studiare gli integrali doppi e tripli), il volume di questo solidi sarà costituito dalla somma di tutti i vari cerchi che si sono formati per rotazione di ciascun punto del grafico. Ognuno di questi cerchi avrà all'ascissa x area pari a $\pi f^2(x)$, quindi il volume del solido sarà dato da:

$$V = \pi \int_a^b f^2(x)\,\mathrm{d}x$$

Esempio 7.11.1 *Calcolare il volume della sfera di raggio* 1*.*

La sfera può essere considerato un solido di rotazione, per la precisione dato dalla rotazione intorno all'asse x *del semicerchio di raggio* 1*. Questo semicerchio avrà come equazione analitica:*

$$f(x) = \sqrt{1-x^2}$$

Definita tra -1 *ed* 1*; quindi la formula da utilizzare per calcolarne il volume sarà:*

$$\pi \int_{-1}^{1} \left(\sqrt{1-x^2}\right)^2 \mathrm{d}x = \pi \int_{-1}^{1} (1-x^2)\,\mathrm{d}x = \pi\left[x - \frac{x^3}{3}\right]_{-1}^{1} = \frac{4}{3}\pi$$

E questo risultato effettivamente coincide con la formula che già si conosce per il calcolo del volume di una sfera.

Capitolo 8

Vettori, matrici e sistemi lineari

8.1 Vettori

Prima di poter passare all'esame delle funzioni in più variabili, occorre studiare gli spazi di tipo $\mathbb{R}^n$. Questi insiemi numerici sono costituiti da n-uple ordinate di numeri di $\mathbb{R}$, chiamati *vettori*. Il termine *ordinate* significa, molto semplicemente, che, ad esempio, il vettore $(1, 2, 3)$ è diverso dal vettore $(1, 3, 2)$: l'ordine è importante.

Esiste anche una facile interpretazione geometrica dei vettori. Prendendo i casi in cui $n = 2$ ed $n = 3$, i vettori possono essere rappresentati da punti degli spazi cartesiani bidimensionali o tridimensionali.

In generale, un vettore di componenti x_1, x_2, ..., x_n si indica con la scrittura:

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

Una variabile di tipo vettore si può indicare in vari modi:

$$\vec{x}, \quad \overline{x}, \quad \boldsymbol{x}$$

In questo testo sarà utilizzata la prima notazione.

Se prendiamo due vettori $\vec{x} = (x_1, x_2, \ldots, x_n)$ e $\vec{x}' = (x'_1, x'_2, \ldots, x'_n)$, allora si pone per definizione:

$$\vec{x} + \vec{x}' \stackrel{def}{=} (x_1 + x'_1, x_2 + x'_2, \ldots, x_n + x'_n)$$

La somma così definita risulta sempre essere commutativa, associativa, possiede elemento neutro ($\vec{0} = (0, 0, \ldots, 0)$) ed ogni vettore possiede un inverso ($-\vec{x} = (-x_1, -x_2, \ldots, -x_n)$). Si scopre facilmente che il prodotto tra vettori non è definibile come prodotto dei loro singoli componenti – per lo meno non senza perdere le proprietà che già conosciamo del prodotto. Tuttavia, è facilmente definibile il prodotto per uno scalare:

$$\lambda\vec{x} = (\lambda x_1, \lambda x_2, \ldots, \lambda x_n)$$

Si può inoltre definire il concetto di *distanza* tra due punti, o vettori, di $\mathbb{R}^n$ estendendo semplicemente il concetto di distanza euclidea:

$$\mathrm{dist}_{\vec{x}\to\vec{y}} \stackrel{def}{=} \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \ldots + (x_n - y_n)^2}$$

Anche in questo caso, la definizione di distanza rispetta tutte le proprietà già studiate in $\mathbb{R}$. Inoltre, se prendiamo il caso $n = 1$, abbiamo $\vec{x} = (x_1)$ e $\vec{y} = (y_1)$, e la distanza diventa:

$$\mathrm{dist}_{\vec{x}\to\vec{y}} = \sqrt{(x_1 - y_1)^2} = |x_1 - y_1|$$

Che è effettivamente la definizione già usata in precedenza.

Possiamo quindi anche dire che:

$$|\vec{x}| \stackrel{def}{=} \mathrm{dist}_{\vec{x}\to\vec{0}}$$

e che

$$|\vec{x} - \vec{y}| \stackrel{def}{=} \mathrm{dist}_{\vec{x}\to\vec{y}}$$

Bisogna ricordare, come già fatto notare nello studio dei numeri complessi, che in un $\mathbb{R}^n$ generico *non esiste* alcuna relazione d'ordine che rispetti le proprietà che si possono trovare in $\mathbb{R}$.

8.1.1 Dipendenza lineare

Prendiamo k vettori di $\mathbb{R}^n$:

$$\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$$

e k scalari

$$\lambda_1, \lambda_2, \ldots, \lambda_k$$

A questo punto se considero la somma

$$\lambda_1\vec{v}_1 + \lambda_2\vec{v}_2 + \ldots + \lambda_k\vec{v}_k$$

Ho costruito una *combinazione lineare* di k vettori.

Un insieme di vettori $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$ si dicono *linearmente dipendenti* se esiste una combinazione lineare che valga 0 senza che tutti i suoi coefficienti siano nulli. In caso contrario si dicono *lineamente indipendenti*.

Questo in pratica significa che non è possibile scrivere un vettore come combinazione lineare dei restanti.

Si verifica facilmente che questa ultima condizione in $\mathbb{R}^2$ coincide semplicemente con l'imporre che i due vettori non siano l'uno multiplo dell'altro.

Un importante teorema afferma che:

Teorema 8.1.1 *In $\mathbb{R}^n$, k vettori $\vec{v}_k$ sono dipendenti se $k > n$ – il massimo numero di vettori indipendenti è n.*

Questo ad esempio significa che 3 vettori in $\mathbb{R}^2$ sono obbligatoriamente dipendenti, ed al massimo si possono avere 2 vettori indipendenti.

n vettori indipendenti di $\mathbb{R}^n$ si dice che formano una *base* di $\mathbb{R}^n$. Ad esempio, una classica base di $\mathbb{R}^2$ è la coppia

$$\begin{aligned}\vec{i} &= (1,0)\\ \vec{j} &= (0,1)\end{aligned}$$

Dimostriamo che $\vec{i}$ e $\vec{j}$ sono indipendenti. Infatti, $\lambda\vec{i} + \mu\vec{j} = (\lambda, \mu)$. Ora, dire che sono linearmente indipendenti significa che il fatto che una loro generica combinazione lineare è nulla implica che i coefficienti della combinazione lineare sono nulli anch'essi. Ed in effetti, in modo alquanto ovvio:

$$(\lambda, \mu) = \vec{0} \Longrightarrow \lambda = 0 \wedge \mu = 0$$

Con un metodo del genere di può dimostrare che una qualunque n-upla di vettori:

$$\begin{aligned}\vec{i}_1 &= \overbrace{(1,0,0,\ldots,0,0)}^{\text{n componenti}}\\ \vec{i}_2 &= (0,1,0,\ldots,0,0)\\ &\ldots\\ \vec{i}_{n-1} &= (0,0,0,\ldots,1,0)\\ \vec{i}_n &= (0,0,0,\ldots,0,1)\end{aligned}$$

In uno spazio $\mathbb{R}^n$ è una base per questo spazio.

Per ritornare al nostro esempio, facilmente estendibile a spazi con dimensione maggiore, si può notare che qualunque altro vettore di $\mathbb{R}^2$ può essere scritto come combinazione lineare di $\vec{i}$ e $\vec{j}$ – infatti, per il teorema 8.1.1 una terna di vettori è sempre linearmente dipendente.

La coppia $\vec{i}$, $\vec{j}$ (o la n-upla $\vec{i}_1$, $\vec{i}_2$, …, $\vec{i}_n$) viene detta base *standard* o *canonica* di $\mathbb{R}^2$ (o di $\mathbb{R}^n$).

A parte la base canonica di uno spazio, ne esistono infinite altre. Ad esempio, sempre in $\mathbb{R}^2$, prendiamo la coppia:

$$\begin{aligned}\vec{v}_1 &= (1,-1)\\ \vec{v}_2 &= (1,1)\end{aligned}$$

Anch'essa è una base del nostr spazio; infatti:

$$\lambda\vec{v}_1 + \mu\vec{v}_2 = (\lambda + \mu, -\lambda + \mu)$$

$$\begin{cases}\lambda + \mu = 0\\ -\lambda + \mu = 0\end{cases}$$

$$\begin{cases}2\mu = 0\\ \lambda = \mu\end{cases}$$

$$\begin{cases}\lambda = 0\\ \mu = 0\end{cases}$$

Ecco quindi che abbiamo dimostrato l'indipendenza di questa coppia di vettori. Quindi, dovremmo essere in grado di rappresentare un qualunque vettore $\vec{x} = (x_1, x_2)$ sotto forma di combinazione lineare di questa base, ovvero $\vec{x} = \lambda\vec{v}_1 + \mu\vec{v}_2$. Da qui si ricava:

$$\begin{cases}x_1 = \lambda + \mu\\ x_2 = -\lambda + \mu\end{cases}$$

Da cui le soluzioni:

$$\begin{cases}\lambda = \dfrac{x_1 + x_2}{2}\\[2ex] \mu = \dfrac{x_1 - x_2}{2}\end{cases}$$

Osserviamo inoltre questo fatto: se la coppia di vettori

$$\vec{u} = (a, b)$$
$$\vec{v} = (c, d)$$

Sono dipendenti, questo significa che $\vec{u} = \lambda\vec{v}$, ovvero che:

$$a = \lambda c$$
$$b = \lambda d$$

Ovvero

$$\frac{a}{c} = \frac{b}{d}$$
$$\frac{ad - bc}{cd} = 0$$

Ma quest'ultima condizione equivale a:

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = 0$$

Vedremo che questa osservazione si potrà applicare anche a spazi a più di due dimensioni.

Viene chiamato *versore* un vettore di lunghezza unitaria. Tutti i vettori delle basi canoniche di uno spazio sono quindi versori.

8.2 Matrici

Una *matrice p per q* è, molto semplicemente, una "tabella" a p righe e q colonne (di numeri reali, nel nostro studio); si ha quindi che $p, q \in \mathbb{N}$, $p, q \geq 1$. Se $p = q$, la matrice si dice *quadrata*.

Se prendiamo ad esempio una matrice $A = (a_{ij})$ intendiamo dire che è una matrice che ha per elementi l'insieme dei numeri reali a_{ij}, dove l'indice i indica la riga e j la colonna. Quindi, per fare un esempio, l'elemento a_{14} indica l'elemento della matrice presente alla prima riga e quarta colonna.

Per le matrici è facile definire in modo piuttosto naturale le operazioni di somma, moltiplicazione per uno scalare, e gli elementi opposto e neutro:

$$A_{p\times q} + B_{p\times q} = C_{p\times q} \quad \Longrightarrow \quad c_{ij} \stackrel{def}{=} a_{ij} + b_{ij}$$
$$\lambda A = C, \lambda \in \mathbb{R} \quad \Longrightarrow \quad c_{ij} = \lambda a_{ij}$$

L'elemento neutro della somma, come è facile intuire, è la matrice contenente solo degli zeri al suo interno. L'opposto sarà quindi:

$$-A = C: \quad c_{ij} = -a_{ij}$$

8.2.1 Prodotto "righe per colonne"

Il prodotto che si definisce per le matrici è il cosidetto *prodotto righe per colonne*. Inanzitutto, la condizione necessaria e sufficiente perchè sia possibile moltiplicare due matrici tra di loro è che il numero di colonne della prima sia pari al numero di righe della seconda. Il risultato avrà il numero di righe della prima ed il numero di colonne della seconda. Scritto in simboli:

$$A_{p\times q} \cdot B_{q\times r} = C_{p\times r}$$

La definizione che si da per questo prodotto è la seguente:

$$c_{ij} = \sum_{v=1}^{q} a_{iv} \cdot b_{vj}$$

In realtà è molto più semplice di quanto possa sembrare. Ad esempio, se dovessimo calcolare l'elemento c_{25} dovremmo considerare la seconda riga della prima matrice e la quinta colonna della seconda matrice. Ora, esse avranno lo stesso numero di elementi (basta ricordare la condizione necessaria per compiere il prodotto tra matrici): a questo punto li moltiplichiamo a due a due, e sommiamo i risultati ottenuti. Ecco trovato l'elemento c_{25}!.

Esempio 8.2.1 *Calcolare*

$$A \cdot B$$

Sapendo che

$$A = \begin{bmatrix} 1 & -\frac{1}{2} & 0 \\ \frac{1}{7} & -2 & \sqrt{2} \end{bmatrix}$$

$$B = \begin{bmatrix} 0 & 1 \\ \frac{1}{3} & \frac{1}{5} \\ -2 & -\sqrt{2} \end{bmatrix}$$

Dato che A ha 3 colonne e B 3 righe, possiamo calcolare il prodotto. Come da definizione:

$$c_{11} = \sum_{v=1}^{3} a_{1v} \cdot b_{v1} = a_{11} \cdot b_{11} + a_{12} \cdot b_{21} + a_{31} \cdot b_{13} = 0 - \frac{1}{6} + 0 = -\frac{1}{6}$$

$$c_{12} = 1 \cdot 1 + \left(-\frac{1}{2}\right)\left(-\frac{1}{5}\right) + 0 \cdot \sqrt{2} = 1 + \frac{1}{10} = \frac{11}{10}$$

$$c_{21} = 0 \cdot \frac{1}{7} + \frac{1}{3} \cdot (-2) + (-2) \cdot \sqrt{2} = -\frac{2}{3} - 2\sqrt{2}$$

Analoghi calcoli per c_{22}. La matrice che così si ottiene è la seguente:

$$C = \begin{bmatrix} -\frac{1}{6} & \frac{11}{10} \\ -\frac{2}{3} - 2\sqrt{2} & -\frac{79}{35} \end{bmatrix}$$

Bisogna ricordare alcuni fatti fondamentali sul prodotto di matrici. Inanzitutto, in linea generale:

$$A \times B \neq B \times A$$

Infatti:

- Se si può calcolare $A \times B$ non è detto che si possa calcolare $B \times A$.
- Anche se si può fare, a meno che A e B non siano matrici quadrate della stessa dimensione, le dimensioni di $A \times B$ saranno diverse da quelle di $B \times A$.
- Infine, anche quando le due matrici risultato hanno la stessa dimensione, basta leggere la formula che le definisce per convincersi che non è detto che siano uguali:

$$c_{ij} = \sum_{v=1}^{q} a_{iv} \cdot b_{vj}$$

$$c'_{ij} = \sum_{v=1}^{q} b_{iv} \cdot a_{vj}$$

Esempio 8.2.2 *Date:*

$$A = \begin{bmatrix} 1 & -1/2 \\ 2 & 1/3 \end{bmatrix}$$

$$B = \begin{bmatrix} 0 & 7 \\ 0 & 0 \end{bmatrix}$$

Verificare che $A \times B \neq B \times A$.

Infatti:

$$A \times B = \begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix}$$

Mentre:

$$B \times A = \begin{bmatrix} 2 & 1/3 \\ 0 & 0 \end{bmatrix}$$

Il prodotto tra matrici gode anche di alcune proprietà:

$$A \times (B + C) = A \times B + A \times C \qquad \text{(Proprietà distributiva)}$$
$$A \times (B \times C) = (A \times B) \times C \qquad \text{(Proprietà associativa)}$$

Anche in questo caso è possibile trovare l'elemento neutro della moltiplicazione, che è la cosiddetta *matrice identità*:

$$I_p = \underbrace{\left.\begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}\right\}}_{\text{p colonne}} \text{p righe}$$

Infatti è facile verificare che:

$$A_{p\times q} \times I_q = A_{p\times q}$$

E

$$I_p \times A_{p\times q} = A_{p\times q}$$

Quindi, l'unica operazione che rimane da definire è *l'inversa* di una matrice.
Una matrice quadrata $A_{p\times p}$ si dice *invertibile* se

$$\exists B_{p\times p} \mid \begin{cases} AB = I_p \\ BA = I_p \end{cases}$$

Questa matrice B deve essere obbligatoriamente *unica*. Infatti, se ciò non fosse, ovvero se:

$$\exists B' \neq B \ : \ AB' = B'A = I_p$$

Avremmo anche che:

$$AB = I_p$$

Per la definizione data prima. Moltiplicando entrambi i membri per B':

$$B'AB = B'I_p$$

E quindi

$$(B'A)B = B'$$

Per ipotesi:

$$I_pB = B'$$

Ovvero

$$B = B'$$

Che, per ipotesi, non può essere vero. Quindi, per assurdo, abbiamo dimostrato che la matrice inversa è unica.

Se prendiamo il caso $p = 1$, possiamo identificare le matrici $A_{1\times 1}$ con il loro singolo elemento a; in questo caso la loro inversa sarà $B_{1\times 1} = b$. Quindi avremo che $AB = ab$, come è facile verificare, e da qui otteniamo che l'inversa di A sarà $A^{-1} = (1/a)$ come ci aspettavamo.

Adesso però abbiamo trovato un problema interessante, ovvero la condizione necessaria e sufficiente per l'invertibilità di una matrice. In realtà esiste una famiglia di problemi che, come vedremo, saranno collegati tra loro:

- Quando la matrice A è invertibile?
- Presi k vettori in $\mathbb{R}^n$, $k \leq n$, quando sono linearmente indipendenti?
- Quando il *sistema lineare* ad n equazioni ed n incognite è *determinato*?

Rimane solo da esporre l'ultimo di questi problemi prima di trarre le conclusioni.

8.3 Sistemi lineari

Un *sistema lineare* a p equazioni e q incognite è un sistema nella forma:

$$\begin{cases} \sum_{j=1}^{q} a_{1j}x_j = y_1 \\ \sum_{j=1}^{q} a_{2j}x_j = y_2 \\ \dots \\ \sum_{j=1}^{q} a_{pj}x_j = y_p \end{cases}$$

Ad esempio il seguente è un sistema lineare a 4 incognite e 2 equazioni:

$$\begin{cases} 2x_1 - \sqrt{3}x_2 + 5x_3 + 3x_4 = 1/4 \\ 6x_1 - 4x_2 + 2x_3 + 6x_4 = \ln 5 \end{cases}$$

In questi casi si è solito prendere, per definire il sistema, alcune matrici che lo definiscono; la prima è la cosiddetta *matrice dei coefficienti*, che in questo caso risulterebbe essere:

$$A = \begin{bmatrix} 2 & -\sqrt{3} & 5 & 3 \\ 6 & -4 & 2 & 6 \end{bmatrix}$$

Ovvero è la matrice costituita da tutti i coefficienti delle incognite. Nella forma generale riportata all'inizio, questa matrice sarebbe molto semplicemente:

$$A = (a_{ij})$$

Quindi c'è la *matrice delle incognite*. Questa matrice nel nostro caso risulta essere:

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

O, nel caso generale:

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_q \end{bmatrix}$$

Questa matrice può essere vista anche come un vettore, detto *vettore colonna*, poichè è rappresentato con una colonna di una matrice. È alquanto intuitivo capire cosa sia un vettore riga...

Infine si prende la *matrice dei termini noti*, che nel nostro caso sarebbe:

$$Y = \begin{bmatrix} 1/4 \\ \ln 5 \end{bmatrix}$$

O, nel caso generale:

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{bmatrix}$$

Anche questa matrice è quindi un vettore colonna.

Perchè utilizzare questa rappresentazione? Perchè in realà l'intero sistema, utilizzando questa notazione, può essere riassunto nella seguente forma:

$$A \times X = Y$$

E ce ne si può convincere facilmente sviluppando i calcoli utilizzando la regola del prodotto tra matrici che si è visto precedentemente. Ovviamente, questa considerazione è vera anche per i casi generali considerati. D'altronde, come verifica superficiale, basta vedere che:

$$A_{p\times q} X_{q\times 1} = Y_{p\times 1}$$

Il che implica che questo prodotto sia sempre possibile.

In pratica, quindi, quando ci si trova a risolvere un sistema lineare, lo scopo è trovare la matrice delle incognite. Di particolare interesse sono i cosiddetti *sistemi lineari quadrati*.

8.3.1 Sistemi lineari quadrati: teorema di Cramer

Viene detto *sistema quadrato*, in modo alquanto intuitivo, un sistema in cui il numero di equazioni sia pari a quello delle incognite, ovvero in cui la matrice dei coefficienti sia quadrata.

Viene chiamato *sistema omogeneo* un sistema nella forma:

$$AX = 0$$

Ovvero nel quale la matrice dei termini noti sia costituita solo da degli zeri.

Per i sistemi lineari quadrati esiste un importante teorema:

Teorema 8.3.1 (Teorema di Cramer) *Sono equivalenti le seguenti tre affermazioni:*

a) La matrice dei cofficienti A è invertibile.

b) $\forall Y$, il sistema $AX = Y$ ha una ed una sola soluzione.

c) $AX = 0$ ha come unica soluzione $X = 0$.

La sua dimostrazione quindi implica di dimostrare l'equivalenza delle tre proposizioni a, b e c – questo sarà fatto secondo lo schema $a \Rightarrow b$, $b \Rightarrow a$, $b \Rightarrow c$, $c \Rightarrow b$.

$a \Rightarrow b$

A è invertibile, quindi da

$$AX = Y$$

Posso scrivere

$$A^{-1}(AX) = A^{-1}Y$$
$$(A^{-1}A)X = A^{-1}Y$$
$$IX = A^{-1}Y$$
$$X = A^{-1}Y$$

Ecco che abbiamo trovato il valore di X, che quindi risulta unico. Verificare che questa è effettivamente la soluzione del sistema, non è difficile:

$$A(A^{-1}Y) = Y$$
$$IY = Y$$
$$Y = Y$$

Non solo quindi abbiamo dimostrato l'unicità della soluzione di questo sistema, ma ne abbiamo anche trovato la soluzione, che, come si può vedere, è legata alla ricerca dell'inversa di una matrice, come avevamo già preannunciato al paragrafo precedente.

$b \Rightarrow a$

Prendiamo la matrice, o vettore colonna:

$$Y^{(1)} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Per ipotesi (b):

$$\exists X^{(1)} \mid AX^{(1)} = Y^{(1)}$$

Ora prendo un $Y^{(2)}$ così definito:

$$Y^{(2)} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Come prima, abbiamo che:

$$\exists X^{(2)} \mid AX^{(2)} = Y^{(2)}$$

Proseguendo così, arriviamo a dire che

$$\exists X^{(p)} \mid AX^{(p)} = Y^{(p)}$$

Ora costruiamo la seguente matrice:

$$\left[\begin{array}{c|c|c|c} & & & \\ X^{(1)} & X^{(2)} & \dots & X^{(p)} \\ & & & \end{array}\right]$$

Ovvero una matrice dove ogni colonna è rappresentata da uno dei vettori colonna che abbiamo trovato e dell'esistenza dei quali abbiamo certezza.

Ora, possiamo dire che $AB = I_p$, e che quindi esiste l'inversa della matrice A. Per dimostrare questo fatto, bisognerà dimostrare che $AB = I_p$ (in realà occorrerebbe anche dimostrare che $BA = I_p$, ma questa dimostrazione risulta più complessa e quindi verrà saltata).

Dunque, calcoliamo i valori della prima colonna di AB. I suoi vari elementi, in successione a partire dall'alto, saranno i seguenti:

$$\begin{bmatrix} 1^a \text{ riga di A per la } 1^a \text{ colonna di B} \\ 2^a \text{ riga di A per la } 1^a \text{ colonna di B} \\ \vdots \\ p^a \text{ riga di A per la } 1^a \text{ colonna di B} \end{bmatrix}$$

Ma, osservando come è costruito B, notiamo che questa colonna è proprio $AX^{(1)}$, e quindi è uguale a $Y^{(1)}$.

Analogamente la seconda colonna della matrice risulterà uguale ad $AX^{(2)}$, e così via. Allora, la matrice risultante sarà:

$$\left[\begin{array}{c|c|c|c} & & & \\ Y^{(1)} & Y^{(2)} & \dots & Y^{(p)} \\ & & & \end{array}\right] = \begin{bmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & 0 & 1 \end{bmatrix} = I_p$$

Ecco quindi dimostrato che abbiamo costruito l'inversa della matrice.

$b \Rightarrow c$

Possiamo inanzitutto dire che il sistema possiede una soluzione, che è $X = 0$, infatti:

$$A0 = 0$$

Per qualsiasi A. Inoltre, per l'ipotesi, il sistema $AX = 0$ possiede una ed una sola soluzione.

$c \Rightarrow b$

Scomponiamo la matrice A nel seguente modo:

$$A = \left[\begin{array}{c|c|c|c} & & & \\ V_1 & V_2 & \dots & V_p \\ & & & \end{array}\right]$$

E consideriamo quindi la p-upla di vettori:

$$V_k \equiv \vec{v}_k \in \mathbb{R}^p$$

È possibile dimostrare (e sarà fatto più avanti) che dire che i vettori v_k sono indipendenti tra loro (i v_k sono una base di $\mathbb{R}^p$) è equivalente all'ipotesi data, ovvero che il sistema $AX = 0$ ha come sola soluzione $X = 0$. Questo risultato ci serve per finire la dimostrazione.

Prendiamo quindi un qualsiasi Y:

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{bmatrix}$$

Possiamo quindi prendere, come prima, un vettore equivalente:

$$\vec{y} = (y_1, y_2, \dots, y_p)$$

Ora, se è vero il risultato dato prima, posso esprimere il vettore $\vec{y}$ come combinazione lineare della base v_k:

$$\vec{y} = \alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \dots + \alpha_p\vec{v}_p$$

Dove l'insieme degli α_k è una ben determinata stringa di reali. Ora, calcoliamo il seguente prodotto, e troveremo un risultato molto interessante:

$$A \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \dots \\ \alpha_p \end{bmatrix} = \left[\begin{array}{c|c|c|c} & & & \\ V_1 & V_2 & \dots & V_p \\ & & & \end{array}\right] \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \dots \\ \alpha_p \end{bmatrix}$$

Ora, sviluppando i calcoli, troviamo il valore di questa matrice (indicando con v_{ij} il valore del j-esimo elemento del vettore colonna V_i):

$$\begin{bmatrix} v_{11}\alpha_1 + v_{21}\alpha_2 + \dots + v_{p1}\alpha_p \\ v_{12}\alpha_1 + v_{22}\alpha_2 + \dots + v_{p2}\alpha_p \\ \vdots \\ v_{1p}\alpha_1 + v_{2p}\alpha_2 + \dots + v_{pp}\alpha_p \end{bmatrix}$$

Ora, se osserviamo il valore di ognuna di queste righe, troviamo che la prima riga corrisponde alla prima componente del vettore $\vec{y}$, la seconda riga al secondo componente, e così via, e quindi il risultato di questo prodotto è proprio Y. Allora, il vettore colonne degli α_k è proprio la X che cercavano! E dato che in una base di $\mathbb{R}^p$ esiste una ed una sola combinazione lineare che esprime un qualunque vettore, allora è unico anche il vettore X: abbiamo dimostrato anche questa parte del teorema di Cramer. CVD.

Quindi, dopo tutte queste considerazioni, siamo tornati al punto di partenza: infatti abbiamo trovato anche che il problema di trovare se k vettori sono indipendenti è equivalente a trovare se la matrice associata è invertibile o meno. Ad esempio, dire che i vettori:

$$\vec{v}_1 = (1, \frac{1}{3}, -2)$$

$$\vec{v}_2 = (\sqrt{3}, \frac{1}{5}, 7)$$

$$\vec{v}_3 = (-1, -\frac{1}{12}, 4)$$

sono indipendenti è come dire che la matrice

$$\begin{bmatrix} 1 & \frac{1}{3} & -2 \\ \sqrt{3} & \frac{1}{5} & 7 \\ -1 & -\frac{1}{12} & 4 \end{bmatrix}$$

è invertibile.

8.4 Determinante

Prima di tutto, terminiamo la dimostrazione del teorema di Cramer.

Teorema 8.4.1 *Presi n vettori*

$$\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$$

e la matrice

$$A_{n\times n} = \left[\begin{array}{c|c|c|c} & & & \\ \vec{v}_1 & \vec{v}_2 & \ldots & \vec{v}_n \\ & & & \end{array}\right]$$

Allora le seguenti proposizioni sono equivalenti:

- $AX = 0$ *possiede una sola soluzione* $(X = 0)$.
- $\vec{v}_1 \ldots \vec{v}_n$ *è una base di* $\mathbb{R}^n$.

Dice che la n-upla data di vettori è una base è come dire che l'insieme dei vettori è linearmente indipendente, e quindi che:

$$\lambda_1\vec{v}_1 + \lambda_2\vec{v}_2 + \ldots + \lambda_n\vec{v}_n = \vec{0} \implies \lambda_1, \lambda_2, \ldots, \lambda_n = 0$$

Ora, chiamiamo la combinazione lineare data dei vettori $\vec{V}$, e troviamo facilmente che

$$V_1 = \lambda_1\vec{v}_{11} + \lambda_2\vec{v}_{21} + \ldots + \lambda_n\vec{v}_{n1}$$

$$V_2 = \lambda_1\vec{v}_{12} + \lambda_2\vec{v}_{22} + \ldots + \lambda_n\vec{v}_{n2}$$

E così via. Ma d'altronde, come già avevamo riscontrato nella dimostrazione del teorema di Cramer, abbiamo che:

$$A\begin{bmatrix}\lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n\end{bmatrix} = \begin{bmatrix}V_1 \\ V_2 \\ \vdots \\ V_n\end{bmatrix}$$

Quindi, $\lambda_1\vec{v}_1 + \lambda_2\vec{v}_2 + \ldots + \lambda_n\vec{v}_n = \vec{0}$ biimplica che $AX = 0$ abbia come soluzione $X = 0$.

A questo punto, dobbiamo trovare le seguenti risposte:

- Come si riconosce se la matrice $A_{n\times n}$ è invertibile?
- Come si calcola, in questo caso, A^{-1}?

La risposta a queste domande è contenuto nel calcolo del *determinante*. Il determinante è un valore reale associato ad una matrice quadrata $n \times n$ definito ricorsivamente. Il determinante di una matrice generica

$$A_{p\times q} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1q} \\ a_{21} & a_{22} & \ldots & a_{2q} \\ \vdots & \vdots & \vdots & \vdots \\ a_{p1} & a_{p2} & \ldots & a_{pq} \end{bmatrix}$$

Si indica con

$$\det A$$

O direttamente con:

$$\begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1q} \\ a_{21} & a_{22} & \ldots & a_{2q} \\ \vdots & \vdots & \vdots & \vdots \\ a_{p1} & a_{p2} & \ldots & a_{pq} \end{vmatrix}$$

Chiariamo inanzitutto il significato di *minore complementare* e *complemento algebrico*.

Viene chiamato *minore complementare* di un elemento a_{ij} della matrice A, e si indica con M_{ij}, il determinante della matrice ottenuta rimuovendo dalla matrice A la riga i-esima e la colonna j-esima.

Viene chiamato *complemento algebrico* di un elemento a_{ij} della matrice A il numero $A_{ij} = (-1)^{i+j} M_{ij}$. In pratica:

$$A_{ij} = \begin{cases} M_{ij} \text{ se } i+j \text{ è pari.} \\ -M_{ij} \text{ se } i+j \text{ è dispari.} \end{cases}$$

Il metodo di calcolo del determinante viene allora dato dal seguente teorema:

Teorema 8.4.2 (Teorema di Laplace) *Il valore del determinante della matrice B si calcola nel modo seguente:*

- *Se $n = 1$, allora*
$$\det \begin{bmatrix} a \end{bmatrix} = a$$
- *Se $n > 1$, allora*
$$\det B = \sum_{j=1}^{n} a_{ij} A_{ij}$$
Indipendentemente dalla i scelta. Un altro metodo di calcolo è:
$$\det B = \sum_{i=1}^{n} a_{ij} A_{ij}$$
Indipendentemente dalla j scelta. In pratica, si va per colonne invece che per righe.

In realtà, in aggiunta a questo teorema si aggiunge che:

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

Risultato che è già contenuto nel teorema dato, ma che, per il frequente utilizzo pratico, conviene tenere a mente.

Si noti che la definizione data è effettivamente ricorsiva, poiché per calcolare il determinante di una matrice $n \times n$, occorre calcolare n complementi algebrici, i quali si calcolano a partire dai minori complementari, che si calcolano attraverso i determinanti di una matrice $(n-1) \times (n-1)$.

Nonostante la complessità della definizione data, in realtà il suo calcolo, per quanto laborioso in termini di conti da eseguire, è concettualmente semplice. Inanzitutto, osserviamo alcune proprietà dei determinanti:

- Se in A è presente una riga o colonna di zeri, allora $\det A = 0$.
- Se A è triangolare (tutti i valori al di sopra o al di sotto della diagonale principale sono nulli) o diagonale (tutti i valori eccetto quelli della diagonale principale sono nulli) allora $\det A = a_{11} \cdot a_{22} \cdot \ldots \cdot a_{nn}$.
- Scambiando due righe o colonne, il segno del determinante cambia.
- Se si moltiplica una riga o colonna per $\lambda \in \mathbb{R}$, il determinante viene moltiplicato per λ.
- $\det(\lambda A) = \lambda^n \det A$, $\lambda \in \mathbb{R}$.
- $\det A = \det A^T$.
- $\det(A + B) \neq \det A + \det B$.
- **Teorema 8.4.3 (Teorema di O. Binet)** $\det(A \times B) = \det A \cdot \det B$
- $\det I = 1$ (I è una matrice diagonale).

Infine, vale il seguente importante teorema che da la risposta alla nostra domanda:

Teorema 8.4.4 *A invertibile* $\iff \det A \neq 0$.

Inoltre abbiamo che, se $B = A^{-1}$, allora:

$$AB = I$$

$$\det(AB) = \det I = 1$$

$$\det A \cdot \det B = 1$$

E quindi arriviamo ad un altro importante risultato:

$$\det A^{-1} = \frac{1}{\det A}$$

Rimane però il problema di come si calcola l'inversa di una matrice con determinante non nullo. La risposta è la facile: l'inversa di una matrice B è la seguente:

$$B^{-1} = \frac{1}{\det B} \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ A_{n1} & A_{n2} & \dots & A_{nn} \end{pmatrix}^T$$

Ovvero, l'inversa di una matrice è il prodotto tra l'inverso del determinate della matrice e la trasposta della matrice che ha per elementi i complementi algebrici della matrice.

Esempio 8.4.1 *Calcolare l'inversa della matrice:*

$$B = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

supponendo che il suo determinante sia non nullo.

Come avevamo già trovato:

$$\det B = ad - bc$$

La matrice dei complementi algebrici risulta essere:

$$\begin{pmatrix} d & -c \\ -b & a \end{pmatrix}$$

Il risultato cercato è quindi:

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

8.5 Cenni di geometria elementare

8.5.1 Prodotto scalare

Dati due vettori $\vec{u}$ e $\vec{v}$, il loro *prodotto scalare* o *interno*, denotato con la notazione $\vec{u} \cdot \vec{v}$ è definito dalla formula seguente:

$$\vec{u} \cdot \vec{v} \stackrel{def}{=} |\vec{u}| \cdot |\vec{v}| \cdot \cos\alpha$$

Dove α è l'angolo che essi formano ($0 \leq \alpha \leq \pi$).

Il prodotto scalare così definito gode di alcune proprietà:

$\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$	(Proprietà commutativa)
$\vec{u} \cdot (\vec{v} + \vec{w}) = \vec{u} \cdot \vec{v} + \vec{u} \cdot \vec{w}$	(Proprietà distributiva)
$\forall t \in \mathbb{R} : (t\vec{u}) \cdot \vec{v} = t(\vec{u} \cdot \vec{v})$	
$\vec{u} \cdot \vec{u} = \lvert\vec{u}\rvert^2$	
$\vec{u} \perp \vec{v} \iff \vec{u} \cdot \vec{v} = 0$	(Condizione di perpendicolarità)

Utilizzando queste proprietà possiamo facilmente calcolare il prodotto scalare di due vettori di $\mathbb{R}^n$.

Inanzitutto consideriamo la sua base canonica $e_1, e_2, \dots, e_n$. Tutti questi vettori saranno ovviamente perpendicolari, infatti il prodotto scalare tra qualunque coppia di questi vettori da risultato nullo a meno che non sia il prodotto di uno di essi per sè stesso, nel qual caso, sempre per le proprietà enunciate, varrà 1.

Presi i vettori

$$\vec{u} = u_1\vec{e}_1 + u_2\vec{e}_2 + \ldots + u_n\vec{e}_n = \sum_{i=1}^{n} u_i\vec{e}_i$$
$$\vec{v} = v_1\vec{e}_1 + v_2\vec{e}_2 + \ldots + v_n\vec{e}_n = \sum_{i=1}^{n} v_i\vec{e}_i$$

scriviamone il prodotto scalare:

$$\vec{u} \cdot \vec{v} = \left(\sum_{i=1}^{n} u_i\vec{e}_i\right) \cdot \left(\sum_{i=1}^{n} v_i\vec{e}_i\right)$$

Utilizzando la proprietà distribuitiva, otteniamo vari prodotti tra scalari e vettori. Ma questi vettori sono solo i vettori della base canonica, i cui risultati già conosciamo. Da qui, otteniamo che solo i prodotti dei corrispondenti componenti rimangono nella somma finale, ovvero il risultato è:

$$\vec{u} \cdot \vec{v} = u_1v_1 + u_2v_2 + \ldots + u_nv_n = \sum_{i=1}^{n} u_iv_i$$

Ovvero: il prodotto scalare tra due vettori è la somma dei prodotti dei loro corrispondenti componenti. In particolare, per $\mathbb{R}^2$ abbiamo:

$$\vec{u} \cdot \vec{v} = u_1 v_1 + u_2 v_2$$

ed in $\mathbb{R}^3$:

$$\vec{u} \cdot \vec{v} = u_1 v_1 + u_2 v_2 + u_3 v_3$$

8.5.2 Prodotto vettoriale

Dati due vettori $\vec{u}$ e $\vec{v}$, il loro prodotto vettoriale, denotato da $\vec{u} \wedge \vec{v}$, è così definito (solo su vettori di $\mathbb{R}^3$):

1. Il modulo di $\vec{u} \wedge \vec{v}$ è dato dalla formula:
$$|\vec{u} \wedge \vec{v}| = |\vec{u}| \cdot |\vec{v}| \sin \alpha$$
Dove come al solito α è l'angolo formato dai due vettori ($0 \leq \alpha \leq \pi$).
2. $\vec{u} \wedge \vec{u}$ è perpendicolare a $\vec{u}$ e $\vec{v}$ (e quindi al piano dei due vettori).
3. $\vec{u}$, $\vec{v}$, $\vec{u} \wedge \vec{v}$ formano, nell'ordine, una terna destrorsa di vettori.

per determinare cos'è una terna destrorsa di vettori, si può usare la cosidetta regola della mano destra: il pollice si mette nella direzione del primo vettore, l'indice in quella del secondo vettore, ed il medio, perpendicolarmente agli altri due, nella direzione dell'ultimo.

Anche il prodotto vettoriale ha varie proprietà:

$$\begin{array}{ll}
\vec{u} \wedge \vec{v} = -\vec{v} \wedge \vec{u} & \text{(Proprietà anticommutativa)} \\
\vec{u} \wedge (\vec{v} + \vec{w}) = \vec{u} \wedge \vec{v} + \vec{u} \wedge \vec{w} & \text{(Proprietà distributiva)} \\
\forall t \in \mathbb{R} : (t\vec{u}) \wedge \vec{v} = t(\vec{u} \wedge \vec{v}) & \\
\vec{u} \wedge \vec{u} = \vec{0} & \\
\vec{u} \parallel \vec{v} \iff \vec{u} \wedge \vec{v} = 0 & \text{(Condizione di parallelismo)}
\end{array}$$

Anche in questo caso grazie alle proprietà del prodotto vettoriale, possiamo determinare il risultato a partire dalle componenti degli operandi. In questo caso, essendo il prodotto vettoriale definito solo in $\mathbb{R}^3$, possiamo direttamente prendere la base canonica $\vec{i}$, $\vec{j}$ e $\vec{k}$, e i due vettori:

$$\begin{aligned}
\vec{u} &= u_1\vec{i} + u_2\vec{j} + u_3\vec{k} \\
\vec{v} &= v_1\vec{i} + v_2\vec{j} + v_3\vec{k}
\end{aligned}$$

Il loro prodotto quindi sarà:

$$(u_1\vec{i} + u_2\vec{j} + u_3\vec{k}) \wedge (v_1\vec{i} + v_2\vec{j} + v_3\vec{k})$$

E utilizzando la proprietà distributiva si ottiene:

$$\begin{array}{l}
u_1v_1\vec{i} \wedge \vec{i} \ + \ u_1v_2\vec{i} \wedge \vec{j} \ + \ u_1v_3\vec{i} \wedge \vec{k} \ + \\
u_2v_1\vec{j} \wedge \vec{i} \ + \ u_2v_2\vec{j} \wedge \vec{j} \ + \ u_2v_3\vec{j} \wedge \vec{k} \ + \\
u_3v_1\vec{k} \wedge \vec{i} \ + \ u_3v_2\vec{k} \wedge \vec{j} \ + \ u_3v_3\vec{k} \wedge \vec{k}
\end{array}$$

In realtà questa formula così complicata contiene solo prodotti tra scalari e vettori, che sappiamo calcolare, ed operazioni di prodotto scalare tra i vettori della base canonica – quindi, una volta che avremo i risultati per questi, potremo arrivare al nostro risultato. D'altronde, come si verifica in fretta, i risultati sono i seguenti:

$$\begin{aligned}
\vec{i} \wedge \vec{j} &= \vec{k} \\
\vec{i} \wedge \vec{k} &= -\vec{j} \\
\vec{j} \wedge \vec{k} &= \vec{i}
\end{aligned}$$

Poi, per mezzo della proprietà anticommutativa, possiamo facilmente ricavare gli altri prodotti:

$$\begin{aligned}
\vec{j} \wedge \vec{i} &= -\vec{k} \\
\vec{k} \wedge \vec{i} &= \vec{j} \\
\vec{k} \wedge \vec{j} &= -\vec{i}
\end{aligned}$$

Ed ora possiamo tornare al nostro complesso risultato e trasformarlo per mezzo dei risultati che abbiamo ottenuto finora:

$$\begin{array}{l}
u_1v_1\vec{0} \ + \ u_1v_2\vec{k} \ + \ u_1v_3(-\vec{j}) \ + \\
u_2v_1(-\vec{k}) \ + \ u_2v_2\vec{0} \ + \ u_2v_3\vec{i} \ + \\
u_3v_1\vec{j} \ + \ u_3v_2(-\vec{i}) \ + \ u_3v_3\vec{0}
\end{array}$$

Ora, raccogliendo e semplificando, otteniamo:

$$(u_2v_3 - u_3v_2)\vec{i} \ - \ (u_1v_3 - u_3v_1)\vec{j} \ + \ (u_1v_2 - u_2v_1)\vec{k}$$

Ovvero, se pensiamo che i coefficienti possono essere riscritti come determinanti di matrici:

$$\begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix} \vec{i} - \begin{vmatrix} u_1 & u_3 \\ v_1 & v_3 \end{vmatrix} \vec{j} + \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix} \vec{k}$$

Questo risultato rimane piuttosto difficile da ricordare, tuttavia lo si può pensare come il determinante della seguente matrice sviluppata sulla prima riga:

$$\begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}$$

Questa non è una vera e propria matrice, in quanto possiede dei vettori al suo interno, ma è in realtà solo un trucco mnemonico, che tuttavia si rivela piuttosto utile.

8.5.3 Rette in $\mathbb{R}^2$ ed $\mathbb{R}^3$

Dato un punto $\vec{p}$ ed una direzione $\vec{v} \neq \vec{0}$, possiamo scrivere la retta definita da essi nel seguente modo:

$$\vec{x} = \vec{p} + t\vec{v}, \quad t \in \mathbb{R}$$

La definizione così data è parametrica: ad ogni valore di t corrisponde un punto della retta. Se poniamo $\vec{p} = (a, b, c)$, $\vec{v} = (\alpha, \beta, \gamma)$ e $\vec{x} = (x, y, z)$ (attenzione a non confondere $\vec{x}$, vettore, con x, scalare e suo componente!), abbiamo le seguenti formule:

$$\begin{cases} x = a + \alpha t \\ y = b + \beta t \\ z = c + \gamma t \end{cases}$$

Che vengono dette *equazioni parametriche della retta*. Nel caso in cui $\alpha, \beta, \gamma \neq 0$, possiamo riscrivere l'espressione precedente nel seguente modo:

$$\frac{x-a}{\alpha} = \frac{x-b}{\beta} = \frac{x-c}{\gamma}$$

Che vengono dette *equazioni cartesiane della retta*. La trasformazione in $\mathbb{R}^2$ è alquanto semplice – in particolare le equazioni cartesiane si possono riscrivere in un'altra forma:

$$y = \frac{\beta}{\alpha}(x-a) + b = \frac{\beta}{\alpha}x + \left(b - a\frac{\beta}{\alpha}\right)$$

Che non è altro che l'equazione della retta come già si conosceva:

$$y = mx + b, \quad m = \frac{\beta}{\alpha}, \quad q = b - a\frac{\beta}{\alpha}$$

8.5.4 Condizione di parallelismo tra le rette

Prese due rette:

$$r: \quad \vec{x} = \vec{p} + t\vec{v}$$

$$s: \quad \vec{x} = \vec{q} + s\vec{u}$$

Esse saranno parallele solo quando i loro vettori di direzione sono uguali, o comunque l'uno multiplo dell'altro, ovvero:

$$\vec{u} = k\vec{v}, k \in \mathbb{R}$$

Oppure, detto in un altro modo, r ed s sono parallele se e solo se $\vec{u}$ e $\vec{v}$ sono dipendenti.

Se le due rette sono parallele, significa che hanno o nessuno o tutti i punti in comune. Se sono coincidenti, significa che hanno almeno un punto in comune, e quindi, riprendendo le stesse due rette, dobbiamo avere che:

$$\exists! t, s : \vec{p} + t\vec{v} = \vec{q} + s'\vec{v}$$

Abbiamo utilizzato $\vec{v}$ ad entrambi i membri poichè sappiamo che sono l'uno multiplo dell'altro.

$$(t - s')\vec{v} = \vec{q} - \vec{p}$$

Quindi, abbiamo che $\vec{q} - \vec{p}$ è un multiplo di $\vec{v}$, ovvero $\vec{q} - \vec{p}$ e $\vec{v}$ (o $\vec{u}$) sono dipendenti. In caso contrario, saranno distinte.

Se prendiamo le rette r ed s non parallele, abbiamo che in $\mathbb{R}^2$ esse hanno per forza un punto in comune, ma non così in $\mathbb{R}^3$!

8.5.5 Piani

Possiamo trovare più modi per definire univocamente un piano π in $\mathbb{R}^3$. Un modo è quello di prendere un punto $\vec{p}$ appartenente al piano ad un vettore $\vec{\nu}$ perpendicolare ad esso. In questo caso la direttrice che va da un qualsiasi punto del piano a $\vec{p}$ $(\vec{x} - \vec{p})$ dovrà essere perpendicolare al vettore perpendicolare al piano, ovvero:

$$(\vec{x} - \vec{p}) \cdot \vec{\nu} = 0$$

Se prendiamo

$$\vec{p} = (a, b, c) = a\vec{i} + b\vec{j} + c\vec{k}$$
$$\vec{\nu} = (\alpha, \beta, \gamma) = \alpha\vec{i} + \beta\vec{j} + \gamma\vec{k}$$
$$\vec{x} = (x, y, z) = x\vec{i} + y\vec{j} + z\vec{k}$$

Possiamo riscrivere la relazione precendente nel modo seguente:

$$[(x-a)\vec{i} + (y-b)\vec{j} + (z-c)\vec{k}] \cdot [\alpha\vec{i} + \beta\vec{j} + \gamma\vec{k}] = 0$$

Ovvero:

$$(x-a)\alpha + (y-b)\beta + (z-c)\gamma = 0$$

Un problema comune è quello di trovare la distanza tra un piano definito in tal modo con i vettori $\vec{p}$ e $\vec{\nu}$, ed il punto $\vec{q}$. Abbiamo inanzitutto che:

$$\vec{q} \notin \pi \Rightarrow (\vec{q} - \vec{p}) \cdot \vec{\nu} \neq 0$$

Poichè se vi appartenesse il problema sarebbe già risolto, ed avrebbe soluzione 0. Cerchiamo inanzitutto la retta passante per $\vec{q}$ ed avente direttrice $\vec{\nu}$, che ovviamente è:

$$\vec{x} = \vec{q} + t\vec{\nu}$$

Ora, l'intersezione tra il piano e la retta sarà quel punto $\vec{x}$ che soddisfa entrambe le equazioni, ovvero per cui:

$$(\vec{q} + t\vec{\nu} - \vec{p}) \cdot \vec{\nu} = 0$$

E semplificando:

$$(\vec{q} - \vec{p}) \cdot \vec{\nu} + t|\vec{\nu}|^2 = 0$$

Risolvendo per t:

$$t = \frac{(\vec{q} - \vec{p}) \cdot \vec{\nu}}{|\vec{\nu}|^2}$$

E quindi, nella formula della retta:

$$\hat{\vec{x}} = \vec{q} + \frac{(\vec{q} - \vec{p}) \cdot \vec{\nu}}{|\vec{\nu}|^2}\vec{\nu}$$

Ora, dobbiamo calcolare la distanza tra quest punto e $\vec{q}$, ovvero:

$$|\hat{\vec{x}} - \vec{q}| = \left|\vec{q} + \frac{(\vec{q} - \vec{p}) \cdot \vec{\nu}}{|\vec{\nu}|^2}\vec{\nu} - \vec{q}\right|$$

Semplificando:

$$\frac{|(\vec{q} - \vec{p}) \cdot \vec{\nu}|}{|\vec{\nu}|^2}|\vec{\nu}| = \frac{|(\vec{q} - \vec{p}) \cdot \vec{\nu}|}{|\vec{\nu}|}$$

Esempio 8.5.1 *Calcolare la distanza tra il piano*

$$\pi: \quad 2(x-1) + \frac{1}{3}y - \frac{1}{2}z = 0$$

E il punto

$$\vec{q} = (1, 1, 1)$$

Inanzitutto troviamo I vettore $\vec{p}$ e $\vec{\nu}$ del piano π. È d'altronde alquanto facile compendere quali sono se riscriviamo il piano nel modo seguente:

$$2(x-1) + \frac{1}{3}(y-0) - \frac{1}{2}(z-0) = 0$$
$$((x, y, z) - (1, 0, 0)) \cdot \left(2, \frac{1}{3}, -\frac{1}{2}\right) = 0$$

Quindi i vettori cercati sono

$$\vec{p} = (1, 0, 0)$$

e

$$\vec{\nu} = \left(2, \frac{1}{3}, -\frac{1}{2}\right)$$

Utilizziamo ora la formula trovata:

$$\vec{q} - \vec{p} = (1,1,1) - (1,0,0) = (0,1,1)$$

$$|(\vec{q} - \vec{p}) \cdot \vec{\nu}| = \left|(0,1,1) \cdot \left(2, \frac{1}{3}, -\frac{1}{2}\right)\right| = \left|\frac{1}{3} - \frac{1}{2}\right| = \frac{1}{6}$$

$$|\vec{\nu}| = \sqrt{2^2 + \left(\frac{1}{3}\right)^2 + \left(-\frac{1}{2}\right)^2} = \sqrt{4 + \frac{1}{9} + \frac{1}{4}} = \frac{\sqrt{157}}{6}$$

$$dist_{\vec{q} \to \pi} = \frac{1}{6} \cdot \frac{6}{\sqrt{157}} = \frac{1}{\sqrt{157}}$$

8.5.6 Condizione di parallelismo tra i piani

Anche in questo caso, la condizione di parallelismo, questa volta tra due piani

$$\pi: \quad (\vec{x} - \vec{p}) \cdot \vec{\nu} = 0$$

e

$$\pi': \quad (\vec{x} - \vec{q}) \cdot \vec{\mu} = 0$$

È data dalla dipendenza lineare di $\vec{\nu}$ e $\vec{\mu}$. Questo fa sì che due piani paralleli si possano sempre riscrivere utilizzando lo stesso vettore perpendicolare al piano ν. Come per le rette, due piani paralleli possono o non aver nessun punto in comune, oppure averli tutti – la domanda diventa allora: quando i due piani paralleli

$$\pi: \quad (\vec{x} - \vec{p}) \cdot \vec{\nu} = 0$$

e

$$\pi': \quad (\vec{x} - \vec{q}) \cdot \vec{\nu} = 0$$

Coincidono? Possiamo inanzitutto riscrivere i due piani nella seguente forma:

$$\pi: \quad \vec{x} \cdot \vec{\nu} = \vec{p} \cdot \vec{\nu}$$

e

$$\pi': \quad \vec{x} \cdot \vec{\nu} = \vec{q} \cdot \vec{\nu}$$

Se i due piani coincidono, vuoi dire che possiedono insiemi di $\vec{x}$ che soddisfano queste equazioni uguali, e quindi avranno anche identici $\vec{x} \cdot \vec{\nu}$, ovvero:

$$\vec{p} \cdot \vec{\nu} = \vec{q} \cdot \vec{\nu}$$

E, portando tutto al primo membro, otteniamo:

$$(\vec{p} - \vec{q}) \cdot \vec{\nu} = 0$$

8.5.7 Equazione della retta intersezione tra due piani

Presi due piani non paralleli

$$\pi: \quad (\vec{x} - \vec{p}) \cdot \vec{\nu} = 0$$

e

$$\pi': \quad (\vec{x} - \vec{q}) \cdot \vec{\mu} = 0$$

(in cui quindi $\vec{\nu}$ e $\vec{\mu}$ sono indipendenti, ovvero l'uno non è multiplo dell'altro). La retta sarà espressa dall'insieme dalle $\vec{x}$ in comune ai due piani, che saranno quindi la soluzione del sistema

$$\begin{cases} (\vec{x} - \vec{p}) \cdot \vec{\nu} = 0 \\ (\vec{x} - \vec{q}) \cdot \vec{\mu} = 0 \end{cases}$$

che può anche essere scritto nella forma

$$\begin{cases} \vec{x} \cdot \vec{\nu} = \vec{p} \cdot \vec{\nu} \\ \vec{x} \cdot \vec{\mu} = \vec{q} \cdot \vec{\mu} \end{cases}$$

E la retta stessa sarà scritta nella forma

$$\vec{x} = \vec{r} + t\vec{\Phi}$$

Sostituendo questa equazione nel primo dei due sistemi, otteniamo:

$$\begin{cases} (\vec{r} + t\vec{\Phi} - \vec{p}) \cdot \vec{\nu} = 0 \\ (\vec{r} + t\vec{\Phi} - \vec{q}) \cdot \vec{\mu} = 0 \end{cases}$$

Ovvero

$$\begin{cases} t\vec{\Phi} \cdot \nu + (\vec{r} - \vec{p}) \cdot \vec{\nu} = 0 \\ t\vec{\Phi} \cdot \mu + (\vec{r} - \vec{q}) \cdot \vec{\mu} = 0 \end{cases}$$

Ma bisogna ricordare che il vettore $\vec{r}$ deve appartenere ad entrambi i piani (basta ricordare il significato geometrico di questo vettore quando si definisce una retta), e quindi soddisferà le equazioni di entrambi i piani quando sostituito alla $\vec{x}$, dando i risultati:

$$(\vec{r} - \vec{p}) \cdot \vec{\nu} = 0$$

e

$$(\vec{r} - \vec{q}) \cdot \vec{\mu} = 0$$

Quindi il sistema che avevamo appena ottenuto si riduce a:

$$\begin{cases} \vec{\Phi} \cdot \vec{\nu} = 0 \\ \vec{\Phi} \cdot \vec{\mu} = 0 \end{cases}$$

Ovvero, se ancora una volta interpretiamo geometricamente il risultato ottenuto, abbiamo che il vettore $\vec{\Phi}$ deve essere perpendicolare ad entrambi i vettori $\vec{\nu}$ e $\vec{\mu}$. Sappiamo anche che, nella definizione di una retta, prendere un certo vettore $\vec{\Phi}$ o un suo qualunque multiplo, non cambia nulla, quindi la condizione che abbiamo trovato ora è sufficiente per determinare il nostro vettore. Ma un modo per esprimere un vettore perpendicolare ad altri due dati lo conosciamo già, ed è quello di usare il prodotto vettoriale; ecco quindi che abbiamo la prima parte della formula cercata:

$$\vec{\Phi} = \vec{\nu} \wedge \vec{\mu}$$

Per quanto riguarda la ricerca del vettore $\vec{r}$, si può invece semplicemente determinare la soluzione utilizzando le equazioni che già prima avevamo scritto:

$$\begin{cases} (\vec{r} - \vec{p}) \cdot \vec{\nu} = 0 \\ (\vec{r} - \vec{q}) \cdot \vec{\mu} = 0 \end{cases}$$

Ma in questo modo abbiamo due equazioni e tre incognite (in quanto il generico vettore che prendiamo $\vec{r}$ si suppone in $\mathbb{R}^{\not 4}$). Tuttavia sappiamo già che $\vec{\nu}$ e $\vec{\mu}$ sono indipendenti, quindi possiamo scrivere $\vec{r} = a\vec{\nu} + b\vec{\mu}$, riducendo il sistema a due incognite (a e b), e quindi:

$$\begin{cases} (a\vec{\nu} + b\vec{\mu}) \cdot \vec{\nu} = \vec{p} \cdot \vec{\nu} \\ (a\vec{\nu} + b\vec{\mu}) \cdot \vec{\mu} = \vec{q} \cdot \vec{\mu} \end{cases}$$

Che diventa

$$\begin{cases} a\vec{\nu} \cdot \vec{\nu} + b\vec{\mu} \cdot \vec{\nu} = \vec{p} \cdot \vec{\nu} \\ a\vec{\nu} \cdot \vec{\mu} + b\vec{\mu} \cdot \vec{\mu} = \vec{q} \cdot \vec{\mu} \end{cases}$$

Semplificando

$$\begin{cases} a|\vec{\nu}|^2 + b\vec{\mu} \cdot \vec{\nu} = \vec{p} \cdot \vec{\nu} \\ a\vec{\nu} \cdot \vec{\mu} + b|\vec{\mu}|^2 = \vec{q} \cdot \vec{\mu} \end{cases}$$

Siamo quindi arrivati ad un sistema lineare quadrato; la sua matrice dei coefficienti è

$$A = \begin{bmatrix} |\vec{\nu}|^2 & \vec{\mu} \cdot \vec{\nu} \\ \vec{\mu} \cdot \vec{\nu} & |\vec{\mu}|^2 \end{bmatrix}$$

Il cui determinante è:

$$\det A = |\vec{\nu}|^2 |\vec{\mu}|^2 - (\vec{\mu} \cdot \vec{\nu})^2$$

Ora, verifichiamo che esso sia sempre diverso da zero e che quindi il sistema dato, per il teorema di Cramer, abbia sempre soluzione. Inanzitutto abbiamo che

$$(\vec{\mu} \cdot \vec{\nu})^2 = (|\vec{\mu}||\vec{\nu}| \cos\alpha)^2 = |\vec{\nu}|^2 |\vec{\mu}|^2 \cos^2\alpha \leq |\vec{\nu}|^2 |\vec{\mu}|^2$$

E quindi il determinante in questione è sempre maggiore di zero se il coseno dell'angolo tra i due vettori non è 1. Ma questo avviene solo se $\alpha = 0$ o $\alpha = \pi$, ed in questi due casi i vettori $\vec{\nu}$ e $\vec{\mu}$ sarebbero dipendenti, cosa che abbiamo detto non avvenire. Quindi, abbiamo che sempre $\det A \neq 0$. Il sistema dato dunque ha sempre soluzione.

8.5.8 Piano per tre punti

Bisogna ricordare inanzitutto che è condizione necessaria il fatto che i tre punti presi in esame ($\vec{q_1}$, $\vec{q_2}$, $\vec{q_3}$) non siano *collineari*.

Detto ciò, possiamo ricondurre questo problema ad uno già risolto. Infatti, abbiamo che un punto appartenente al piano è sicuramente $\vec{q_1}$, ed una retta perpendicolare al piano è la retta perpendicolare ai vettori

$$\vec{q_2} - \vec{q_1}$$

e

$$\vec{q_3} - \vec{q_1}$$

Ma un vettore perpendicolare a questi due è

$$(\vec{q_3} - \vec{q_1}) \wedge (\vec{q_2} - \vec{q_1})$$

E quindi la formula del piano cercato è la seguente:

$$(\vec{x} - \vec{q_1}) \cdot [(\vec{q_3} - \vec{q_1}) \wedge (\vec{q_2} - \vec{q_1})] = 0$$

8.6 Matrici e trasformazioni lineari

Le *trasformazioni lineari* sono una particolare famiglia di funzioni vettoriali a valori vettoriali, ovvero funzioni (o *applicazioni*), che prendono dei vettori come variabile indipendente e restituiscono sempre vettori come variabile dipendente.

Genericamente quindi la trasformazione lineare T sarà della forma:

$$T : \mathbb{R}^q \to \mathbb{R}^p, \quad q, p \in \mathbb{N} \geq 1$$

Vengono dette trasformazioni lineari quelle funzioni che possiedono le seguente proprietà:

1. $T(\vec{u} + \vec{v}) = T(\vec{u}) + T(\vec{v})$
2. $T(\lambda\vec{u}) = \lambda T(\vec{u}), \lambda \in \mathbb{R}$

Ovvero, scritto in un'unica formula:

$$T(\lambda\vec{u} + \gamma\vec{v}) = \lambda T(\vec{u}) + \gamma T(\vec{v})$$

Prendiamo ora il caso più semplice, ovvero quello in cui $p = q = 1$; in questo caso quindi ci troviamo di fronte ad una funzione reale a valori reali:

$$T : \mathbb{R} \to \mathbb{R}$$

Le condizioni ora esposte diventano quindi:

$$f(a + b) = f(a) + f(b)$$

e

$$f(\lambda a) = \lambda f(a)$$

Ora, un qualsiasi vettore $\vec{x}$ in uno spazio unidimensionale può essere rappresentato da un singolo numero: (x). Un'altra rappresentazione è la seguente: $\vec{x} = x \cdot \vec{1}$, dove $\vec{1}$ è il vettore unitario unidimensionale, ovvero $\vec{1} = (1)$. Quindi, otteniamo:

$$f(\vec{x}) = f(x \cdot \vec{1})$$

Ma x è un valore reale, e quindi possiamo utilizzare la seconda proprietà che abbiamo imposto e dire che:

$$f(x \cdot \vec{1}) = xf(\vec{1})$$

Ora, se poniamo $\alpha = f(\vec{1})$ (ricordiamoci che teoricamente la funzione restituisce un vettore, ma essendo esso unidimensionale possiamo ancora una volta considerarlo alla stregua di un numero reale), abbiamo trovato che:

$$f(x) = \alpha x$$

Funzione che rispetta anche la prima proprietà:

$$f(a + b) = \alpha(a + b) = \alpha a + \alpha b = f(a) + f(b)$$

Ecco quindi che, nel caso unidimensionale, le uniche trasformazioni lineari sono le funzioni che rappresentano rette passanti per l'origine – le condizioni imposte sono quindi molto restrittive!

Ora passiamo al caso $\mathbb{R}^q \to \mathbb{R}^p$. Considero una matrice:

$$A_{p \times q} = (a_{ij})$$

E la trasformazione

$$T_A : \mathbb{R}^q \to \mathbb{R}^p$$

Ed un generico vettore

$$\vec{v} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^q$$

Ora definiamo la T_A dichiarata prima nel seguente modo:

$$T_A : \vec{y} \in \mathbb{R}^p \stackrel{def}{=} A \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_q \end{bmatrix}$$

Dove quindi la variabile indipendente è data dal prodotto tra la matrice A e il vettore colonna $\vec{x}$. Effettivamente, essendo $A_{p\times q} \times X_{q\times 1}$, il risultato sarà di dimensioni $Y_{p\times 1}$, e quindi anche il vettore $\vec{y}$ sarà un vettore colonna, questa volta di dimensione p. Ebbene, le trasformazioni così definite sono l'*unico* tipo di trasformazioni lineari possibili.

Dimostriamo allora che queste sono trasformazioni sono lineari.

Prendiamo la base canonica di $\mathbb{R}^q$ ed indichiamola con

$$\begin{aligned} \vec{e}_1 &= (1, 0, 0, \ldots, 0) \\ \vec{e}_2 &= (0, 1, 0, \ldots, 0) \\ &\ldots \\ \vec{e}_q &= (0, 0, 0, \ldots, q) \end{aligned}$$

Allora possiamo scrivere il vettore $\vec{x}$ nel seguente modo:

$$\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2 + \ldots + x_q\vec{e}_q$$

Quindi

$$T(\vec{x}) = T(x_1\vec{e}_1 + x_2\vec{e}_2 + \ldots + x_q\vec{e}_q) = x_1T(\vec{e}_1) + x_2T(\vec{e}_2) + \ldots + x_qT(\vec{e}_q)$$

Quindi, per conoscere il valore di un generico $T(\vec{x})$ basta conoscere i valori di $T(\vec{e}_1)$, $T(\vec{e}_2)$, $\ldots$, $T(\vec{e}_q)$, ovvero un numero finito di valori. Tutti questi vettori ovviamente apparterranno a $\mathbb{R}^p$. Ora, prendiamo la base canonica di $\mathbb{R}^p$, questa volta, ed indichiamola con:

$$\begin{aligned} \vec{\varepsilon}_1 &= (1, 0, 0, \ldots, 0) \\ \vec{\varepsilon}_2 &= (0, 1, 0, \ldots, 0) \\ &\ldots \\ \vec{\varepsilon}_p &= (0, 0, 0, \ldots, p) \end{aligned}$$

Da notare che, essendo genericamente $p \neq q$, allora $\vec{e}_n \neq \vec{\varepsilon}_n$!

Ora quindi possiamo rappresentare ognuno dei $T(\vec{e}_n)$ come combinazione lineare della base $\vec{\varepsilon}_n$.

$$\begin{aligned} T(\vec{e}_1) &= a_{11}\vec{\varepsilon}_1 + a_{21}\vec{\varepsilon}_2 + \ldots + a_{p1}\vec{\varepsilon}_p \\ T(\vec{e}_2) &= a_{12}\vec{\varepsilon}_1 + a_{22}\vec{\varepsilon}_2 + \ldots + a_{p2}\vec{\varepsilon}_p \\ &\ldots \\ T(\vec{e}_q) &= a_{1q}\vec{\varepsilon}_1 + a_{2q}\vec{\varepsilon}_2 + \ldots + a_{pq}\vec{\varepsilon}_p \end{aligned}$$

Quindi così possiamo definire la matrice

$$A_{p\times q} = (a_{ij})$$

Dalla rappresentazione che ne abbiamo dato, ogni colonna di A contiene il vettore colonna $T(\vec{e}_n)$:

$$\left[\begin{array}{c|c|c|c} & & & \\ T(\vec{e}_1) & T(\vec{e}_2) & \ldots & T(\vec{e}_q) \\ & & & \end{array}\right]$$

E quindi a questo punto basta applicare la regola di moltiplicazione tra matrici per constatare che

$$T(\vec{x}) = A \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_q \end{bmatrix}$$

Esempio 8.6.1 *Determinare se la funzione che associa ad un vettore la sua rotazione intorno all'origine di un angolo θ:*

$$T_\theta : \mathbb{R}^2 \to \mathbb{R}^2$$

è una trasformazione lineare, e in questo caso darne l'espressione matriciale.

Inanzitutto esprimiamo le formule che danno questa trasformazione. Nel nostro caso avremo che $\vec{e}_1 = \vec{\varepsilon}_1 = \vec{i}$ e $\vec{e}_2 = \vec{\varepsilon}_2 = \vec{j}$. Sicuramente la trasformazione data è lineare, infatti è facile verificare che

$$T(\vec{a} + \vec{b}) = T(\vec{a}) + T(\vec{b})$$

e

$$T(\lambda\vec{a}) = \lambda T(\vec{a})$$

per mezzo di considerazioni geometriche. Allora, come avevamo visto precendentemente, per determinare la trasformazione, basta avere $T(\vec{i})$ *e* $T(\vec{j})$*. La rotazione di un angolo* θ *dei due versori si ricava per mezzo di considerazioni trigonometriche, e da questi risultati:*

$$T(\vec{i}) = \vec{i}\cos\theta + \vec{j}\sin\theta$$

$$T(\vec{j}) = -\vec{i}\sin\theta + \vec{j}\cos\theta$$

La matrice A sarà quindi data da queste due trasformazioni poste come colonne:

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

Ecco allora che la trasformazione è esprimibile con la formula:

$$T_\theta(\vec{x}) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} \qquad \vec{x} = a\vec{i} + b\vec{j}$$

Esempio 8.6.2 *Determinare la trasformazione lineare che associa ad ogni vettore il suo simmetrico rispetto all'asse delle x.*

Anche in questo caso, il nostro scopo è di trovare $T(\vec{i})$ *e* $T(\vec{j})$*, trovandoci sempre nel caso di una trasformazione da* $\mathbb{R}^2$ *ad* $\mathbb{R}^2$*. In questo caso si ricavano anche più facilmente:*

$$T(\vec{i}) = \vec{i}$$

$$T(\vec{j}) = -\vec{j}$$

E quindi troviamo:

$$T(a\vec{i} + b\vec{j}) = a\vec{i} - b\vec{j}$$

O anche

$$T(a\vec{i} + b\vec{j}) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}$$

Esempio 8.6.3 *Determinare la trasformazione lineare che associa ad ogni punto dello spazio, un punto ruotato di* θ *intorno all'asse zeta.*

L'unica differenza in questo caso è che abbiamo una trasformazione da $\mathbb{R}^3$ *ad* $\mathbb{R}^3$*.*
Con il solito procedimento:

$$T(\vec{i}) = (\cos\theta, \sin\theta, 0)$$

$$T(\vec{j}) = (-\sin\theta, \cos\theta, 0)$$

$$T(\vec{k}) = (0, 0, 1)$$

E quindi:

$$T(a\vec{i} + b\vec{j} + c\vec{k}) = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

Nei precedenti esempi abbiamo trovato vari esempi di importanti trasformazioni lineari, tuttavia bisogna ricordare che, nonostante ciò, la maggior parte delle funzioni vettoriali a valori vettoriali sono *non* lineari. Prendiamo ad esempio

$$f : (x, y) \in \mathbb{R}^2 \to (x^2, y + x^2) \in \mathbb{R}^2$$

Questa trasformazione è chiaramente non lineare. Infatti se prendiamo

$$f(\vec{u} = (a, b)) = (a^2, b + a^2)$$

e

$$f(\vec{v} = (c, d)) = (c^2, d + c^2)$$

Ora, se fosse una trasformazione lineare,

$$f(\vec{u} + \vec{v}) = f((a + c, b + d)) = (a^2 + 2ac + c^2, b + d + a^2 + 2ac + c^2)$$

Dovrebbe essere uguale a

$$f(\vec{u}) + f(\vec{v}) = (a^2 + c^2, b + d + a^2 + c^2)$$

Cosa che non avviene. Questo era per ribadire che le trasformazioni lineari sono solo una piccola famiglia delle funzioni vettoriali a valori vettoriali!

Da ricordare il fatto che la composizione di trasformazioni lineari è ancora una trasformazione lineare, ovvero, se abbiamo le trasformazioni lineari

$$T : \mathbb{R}^q \to \mathbb{R}^p$$

e

$$S : \mathbb{R}^p \to \mathbb{R}^s$$

Allora la loro composizione

$$S \circ T$$

È ancora una trasformazione lineare. È facile verificare che la matrice di trasformazione associata, se quella della trasformazione T è $A_{p\times q}$ e quella della trasformazione S è $B_{s\times p}$, è:

$$C_{s\times q} = B \times A$$

Una particolare trasformazione lineare è quella che ha per matrice la matrice identità: in questo caso ogni vettore viene trasformato in sè stesso, e la trasformazione prende il nome di *trasformazione identità*. Essa viene indicata con:

$$Id_n : \mathbb{R}^n \to \mathbb{R}^n$$

Se la trasformazione

$$T : \mathbb{R}^n \to \mathbb{R}^n$$

È biettiva (ovvero iniettiva e suriettiva), allora sarà anche invertibile, e la trasformazione inversa

$$T^{-1} : \mathbb{R}^n \to \mathbb{R}^n$$

Avrà per matrice di trasformazione l'inversa della matrice originale, ovvero

$$A_{T^{-1}} = (A_T)^{-1}$$

Esempio 8.6.4 *Sapendo che la trasformazione che compie la rotazione di un vettore in $\mathbb{R}^2$ ha per matrice*

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

Verificare che effettivamente la matrice della trasformazione inversa (che è una rotazione in senso antiorario) è l'inversa della matrice A.

Inanzitutto, troviamo la matrice della trasformazione inversa. Se essa è una rotazione in senso antiorario, basterà sostituire $-\theta$ a θ:

$$A = \begin{bmatrix} \cos(-\theta) & -\sin(-\theta) \\ \sin(-\theta) & \cos(-\theta) \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

Ora calcoliamo A^{-1} e verifichiamo che coincide con questa matrice.

$$\det A = \cos\theta\cos\theta - (-\sin\theta)\sin\theta = 1$$

Quindi essa è sempre invertibile.

$$A^{-1} = \frac{1}{\det A}\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}^T = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

Ecco che i due risultati portano lo stesso risultato.

Capitolo 9

Funzioni vettoriali a valori reali

9.1 Introduzione

Dopo questo studio compiuto sui vettori, possiamo cominciare a parlare delle funzioni vettoriali a valori reali, ovvero funzioni genericamente del tipo:

$$f : A \subseteq \mathbb{R}^n \to \mathbb{R}$$

Per esse dovremo in pratica riprendere le stesse considerazioni fatte a riguardo delle funzioni reali a valori reali, ovvero studiare i concetti di limite, continuità, derivabilità e integrabilità.

9.2 Limiti e continuità in $\mathbb{R}^n$

Definiamo ora il concetto di limite. In $\mathbb{R}$ avevamo detto che il limite

$$\lim_{x \to x_0} f(x)$$

Era definito se il punto x_0 era *accessibile* dal dominio di f, ovvero se esisteva almeno una successione z_n, dove $z_n \neq x_0$ e $z_n \to x_0$. Inoltre, il valore di questo limite era k se per tutte le possibili successioni, $f(z_n) \to k$.

Come definiamo il concetto di *accessibilità* in $\mathbb{R}^n$? Basta in realtà sostituire a z_n usa successione di vettori $\vec{a}_n$, lasciando inalterata la definizione. Ovviamente, una successione di vettori tenderà ad un valore $\vec{x}_0$ se ognuna delle sue componenti tende alla rispettiva componente di $\vec{x}_0$.

A questo punto, è facile definire il concetto di *continuità* in $\mathbb{R}^n$; e piú precisamente una funzione f è continua nel punto x_0 se:

$$\lim_{\vec{x} \to \vec{x}_0} f(\vec{x}) = f(\vec{x}_0)$$

Le regole di composizione già viste per determinare se una funzione è continua continuano ovviamente a valere.

9.3 Derivate parziali

Per quanto riguarda il concetto di derivabilità invece già cominciano ad apparire i primi problemi. In $\mathbb{R}$ si diceva che una funzione f era derivabile nel punto x_0 se il limite

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

esisteva finito, e il valore della derivata in quel punto era per l'appunto il valore di questo limite. Se portiamo questo definizione ad $\mathbb{R}^n$, il punto $\vec{x}_0$ dovrà appartenere ad $\mathbb{R}^n$, e quindi anche $\vec{h}$ dovrà fare altrettanto, per essere possibile il sommarlo ad $\vec{x}_0$. Tuttavia, ci troviamo con il problema in seguito di dividere il valore ottenuto per $\vec{h}$, e questa operazione non sappiamo come definirla.

La soluzione che si può trovare è di prendere l'incremento solo in una ben precisa "direzione", ovvero prendo uno degli $\vec{e}_j$, che rappresenta il j-esimo vettore della base canonica di $\mathbb{R}^n$, e quindi trasformiamo il limite come segue:

$$\lim_{t \to 0} \frac{f(\vec{x}_0 + t\vec{e}_j) - f(\vec{x}_0)}{t}$$

Abbiamo compiuto l'incremento di $\vec{x}_0$ solo in una precisa direzione: in pratica è come considerare la funzione ad una solo variabile, la x_j.

La funzione si dice *derivabile rispetto ad* x_j se il limite scritto sopra esiste finito, e il suo valore, che si indica con

$$\frac{\partial f}{\partial x_j}(\vec{x}_0)$$

viene chiamato *derivata parziale di* f *rispetto ad* x_j *nel punto* $\vec{x}_0$.

Nel caso migliore, una funzione f presenta n derivate parziali.

Esempio 9.3.1 *Calcolare le derivate parziali di*

$$f(x,y) = (x+y)^3 x^2$$

Come abbiamo detto, calcolare la derivata parziale di una funzione rispetto ad x_j è come considerare la funzione ad una sola variabile, x_j, per l'appunto, e quindi considerare le altre variabili come costanti.

$$\frac{\partial f}{\partial x}(x,y) = 2x(x+y)^3 + 3x^2(x+y)^2 = x(x+y)^2(5x+2y)$$

$$\frac{\partial f}{\partial y}(x,y) = 3x^2(x+y)^2$$

Esempio 9.3.2 *Calcolare le derivate parziali di*

$$f(x,y) = \arctan\frac{y}{x}$$

Diciamo inanzitutto che il dominio di questa funzione è un sottoinsieme di $\mathbb{R}^2$, ovvero

$$D = \{(x,y) \in \mathbb{R}^2 \mid x \neq 0\}$$

$$\frac{\partial f}{\partial x} = \frac{1}{1+\left(\frac{y}{x}\right)^2} y\left(-\frac{1}{x^2}\right) = -\frac{y}{x^2+y^2}$$

$$\frac{\partial f}{\partial y} = \frac{1}{1+\left(\frac{y}{x}\right)^2}\frac{1}{x} = \frac{x}{x^2+y^2}$$

Se ora esaminiamo il dominio di definizione delle derivate parziali, troviamo che esso è in realtà più esteso *di quello della funzione di partenza, e risulta essere:*

$$D' = \mathbb{R}^2 - (0,0)$$

Questa già avverte che spesso la natura delle funzioni a valori vettoriali può essere ben più complessa di quelle delle funzioni reali.

Adesso tentiamo di capire quali delle applicazioni della derivata prima in $\mathbb{R}$ possono essere trasportate in $\mathbb{R}^n$.

9.3.1 Massimi e minimi relativi

Il concetto di massimo e minimo relativo rimangono. Infatti, ad esempio, prendiamo una funzione

$$f : A \subseteq \mathbb{R}^n \to \mathbb{R}$$

che nel punto $\vec{x}_0$ abbia un massimo; questo significa che vale la seguente:

$$\exists \delta > 0 : \forall \vec{x} \in A, |\vec{x} - \vec{x}_0| < \delta : f(\vec{x}) \leq f(\vec{x}_0)$$

Per il minimo cambierà solo il verso della disuguaglianza. In pratica non abbiamo fatto altro che trasportare la definizione data in $\mathbb{R}$, trasformando le variabili in vettori. Questo è possibile poichè abbiamo definito il concetto di valore assoluto per vettori.

Ora, vale la seguente:

Teorema 9.3.1 (Teorema di Fermat) *Data una funzione $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$, con A insieme aperto, se nel punto $\vec{x}_0 \in A$ esiste un massimo o un minimo, tutte le sue derivate parziali saranno nulle in $\vec{x}_0$*

Per ora non verrà spiegato il significato di "insieme aperto". Basti comunque che tutti gli insiemi $\mathbb{R}$, $\mathbb{R}^2$, ..., $\mathbb{R}^n$ sono validi insiemi aperti per questo teorema. La dimostrazione è comunque semplice: infatti, se restringiamo la funzione ad una singola variabile, considerando le altre costanti, il punto $\vec{x}_0$ deve comunque rappresentare un massimo o minimo di questa funzione reale, e quindi la derivata (parziale, fatta rispetto alla coordinata che abbiamo mantenuto) sarà nulla.

Ovviamente, il contrario, così come non era vero in $\mathbb{R}$, non sarà vero neanche in $\mathbb{R}^n$, ovvero non tutti i punti dove le derivate parziali si annullano sono massimi o minimi.

In $\mathbb{R}$ per determinare quali tra questi punti fossero effettivamente massimi o minimi utilizzavamo lo studio della derivata prima in un intorno del punto, poichè dove sapevamo che, dove la derivata prima era positiva la funzione era monotona crescente, dove invece era negativa essa era monotona decrescente. Ma, come abbiamo già detto, in $\mathbb{R}^n$ non esiste più una relazione di ordine, e quindi neanche un concetto di funzione "crescente" o "decrescente".

Tuttavia, esisteva anche un altro sistema valido, ed era quello di studiare la derivata seconda. Tuttavia in $\mathbb{R}^n$, se è vero che esistono n derivate prime, allora per ognuna di queste esistono n derivate seconde, per un totale di n^2 derivate seconde.

Prima di esaminare a fondo questo concetto, passiamo ad un altro degli utilizzi della derivata prima.

9.3.2 Piani tangenti

Così come in $\mathbb{R}$ potevamo utilizzare la derivata prima di una funzione per determinarne la retta tangente ad essa in un punto. Così, è ragionevole pensare di poter utilizzare le derivate parziali prime di una funzione (che ora consideriamo in $\mathbb{R}^2$) per determinare il *piano* tangente ad essa in ogni suo punto.

Se cerchiamo il piano tangente alla funzione

$$f : \mathbb{R}^{\not 2} \to \mathbb{R}$$

nel punto $\vec{x}_0 = (x_0, y_0)$, sappiamo inanzitutto che esso passa per il punto

$$\vec{p} = (x_0, y_0, f(x_0, y_0))$$

Ora, la generica forma di un piano passate per questo punto sarebbe:

$$a(x - x_0) + b(y - y_0) + c(z - f(x_0, y_0)) = 0$$

Ora rimane quindi solo da determinare a, b e c. Ora, diciamo sin da subito che il valore di questi coefficienti è

$$a = -\frac{\partial f}{\partial x}(x_0, y_0)$$

$$b = -\frac{\partial f}{\partial y}(x_0, y_0)$$

$$c = 1$$

Per determinare se questi sono effettivamente i coefficienti cercati, dobbiamo esaminare prima di tutto come vogliamo definire il piano tangente.

Tornando ad $\mathbb{R}$, avevamo verificato che il concetto di retta tangente era strettamente legato al differenziale, o approssimazione lineare, della funzione stessa. Noi sapevamo infatti che

$$f(x + h) = f(x) + f'(x_0)h + g(h)$$

Dove non solo

$$g(h) \overset{h\to 0}{\to} 0$$

Come era facile aspettarsi con una intuizione geometrica, ma anche

$$\frac{g(h)}{h} \overset{h\to 0}{\to} 0$$

Che è un risultato più restrittivo, poichè la divisione per h, essendo che h tende a 0, porterebbe a pensare che faccia divergere il limite a $\pm\infty$, cosa che non avviene invece.

Ora, il fatto che la retta che passava per i punti $f(x_0)$ e $f(x_0) + f'(x_0)h$ fosse tangente significava proprio questo – che la differenza tra l'ordinata della retta $(f(x_0) + f'(x_0)h)$ e il valore della funzione $(f(x_0 + h))$ era un "o piccolo" di h.

Se cerchiamo l'equivalente in $\mathbb{R}^2$, abbiamo che

$$f(\vec{x}_0 + \vec{h}) - \text{piano}(\vec{x}_0 + \vec{h}) = \Delta$$

e

$$\Delta = o(|\vec{h}|)$$

Abbiamo usato il valore assoluto poichè Δ in effetti è una funzione reale. Ora, utilizzando la formula data per il piano tangente, diamo un valore a questa differenza:

$$\Delta = f(x_0 + h_x, y_0 + h_y) - f(x_0, y_0) - \frac{\partial f}{\partial x}(x_0, y_0)h_x - \frac{\partial f}{\partial y}(x_0, y_0)h_y$$

E quindi il piano dato sarà tangente alla funzione solo se

$$\frac{\Delta}{\sqrt{h_x^2 + h_y^2}} \overset{(h_x,h_y)\to\vec{0}}{\longrightarrow} 0$$

Ora, in $\mathbb{R}$ il semplice fatto che la funzione fosse derivabile assicurava che un'espressione del genere fosse $o(h)$, ovvero che la funzione fosse *differenziabile*. In $\mathbb{R}^n$, con $n > 1$, invece, *ciò non è vero*! Questo si può notare portando un controesempio.

Esempio 9.3.3 *Verificare la funzione*

$$f(x, y) = \begin{cases} \dfrac{xy}{\sqrt{x^2 + y^2}}, & (x, y) \neq (0, 0) \\ 0, & (x, y) = (0, 0) \end{cases}$$

è definita e continua in tutto $\mathbb{R}^2$, ma non possiede piano tangente in $(0,0)$ pur essendo lì derivabile.

Che sia definita in tutto $\mathbb{R}^2$ è ovvio: infatti la prima parte della formula data vale per tutti i punti di $\mathbb{R}^2$ esclusa l'origine, per la quale il denominatore si annulla, ma nell'origine essa è definita con un altro valore, 0, e quindi è definita in tutto il piano.

La dimostrazione della sua continuità è già più complessa, infatti, anche se in tutto $\mathbb{R}^2 - (0,0)$ è ovvio che è continua, non così è in $(0,0)$. Qui quindi si tratta di verificare che

$$\lim_{(x,y)\to(0,0)} f(x) = 0$$

In pratica, dato che dobbiamo considerare la $f(x)$ in tutti i punti di un intorno di $(0,0)$ con $(0,0)$ escluso, la nostra funzione si riduce al solo primo membro:

$$\frac{xy}{\sqrt{x^2+y^2}}$$

Ma la nostra funzione, se sia x che y tendono a 0, diventa una forma indeterminata del tipo $0/0$. In $\mathbb{R}^2$ non si può neanche più utilizzare la regola di De L'Hospital (che effettivamente non vale più), e dobbiamo quindi trovare un altro sistema.

Mostriamo inanzitutto che

$$|xy| \leq \frac{x^2+y^2}{2}$$

Infatti la formula data si può trasformare in

$$2|x||y| \leq |x|^2 + |y|^2$$

$$|x|^2 - 2|x||y| + |y|^2 \geq 0$$

$$(|x| - |y|)^2 \geq 0$$

E dato che quest'ultima proposizione è sempre vera, lo sarà anche quella di partenza. Quindi è anche vero che:

$$0 \leq |xy| \leq \frac{1}{2}(x^2+y^2)$$

Dividendo entrambi i membri per $\sqrt{x^2+y^2}$, che nel dominio da noi considerato è sempre non nullo:

$$0 \leq \frac{|xy|}{\sqrt{x^2+y^2}} \leq \frac{1}{2}\frac{x^2+y^2}{\sqrt{x^2+y^2}}$$

ovvero

$$0 \leq \left|\frac{xy}{\sqrt{x^2+y^2}}\right| \leq \frac{1}{2}\sqrt{x^2+y^2}$$

Ora, per $(x,y) \to (0,0)$, dato che la funzione di destra è continua, avrà per limite $\sqrt{(0^2+0^2)} = 0$, e quindi per il la funzione centrale, essendo limitata tra 0 ed una funzione che anch'essa tende a 0, varrà anch'essa 0.

$$\lim_{(x,y)\to\vec{0}} \left|\frac{xy}{\sqrt{x^2+y^2}}\right| = 0$$

Ma se il valore assoluto della funzione tende a 0, farà così anche la funzione stessa, ovvero:

$$\lim_{(x,y)\to\vec{0}} \frac{xy}{\sqrt{x^2+y^2}} = 0$$

E così abbiamo dimostrato anche la continuità della funzione. Ora dimostriamone la derivabilità calcolandone per l'appunto le derivate parziali nel punto (0,0):

$$\frac{\partial f}{\partial x}(0,0) = \lim_{t\to 0} \frac{f(0+t,0) - f(0,0)}{t} = \lim_{t\to 0} \frac{0}{t} = 0$$

Per simmetria, si può subito concludere che

$$\frac{\partial f}{\partial y}(0,0) = 0$$

Il piano tangente risulterebbe quindi essere, molto semplicemente:

$$z = 0$$

Verifichiamo se è effettivamente così. Se ciò fosse giusto, allora:

$$\lim_{(h_x,h_y)\to\vec{0}} \frac{f(x_0+h_x, y_0+h_y) - piano(x_0,y_0)}{\sqrt{h_x^2+h_y^2}} = \lim_{(h_x,h_y)\to\vec{0}} \frac{f(h_x,h_y)}{\sqrt{h_x^2+h_y^2}} = 0$$

Come prima, la funzione si riduce al suo solo primo membro. La funzione di cui fare il limite diventa poi, cambiando i nomi delle variabili:

$$\frac{f(p,q)}{\sqrt{p^2+q^2}} = \frac{\frac{pq}{\sqrt{p^2+q^2}}}{\sqrt{p^2+q^2}} = \frac{pq}{p^2+q^2}$$

In questo caso il limite si rivela interessante. Infatti, compiendo il limite separatamente nei due assi, ovvero prendendo le successioni $\vec{z}_n = (n, 0)$ *e* $\vec{z}_n = (0, n)$*, in pratica ci si trova a fare dei limiti singoli che tendono effettivamente a* 0*:*

$$\lim_{p\to 0} \frac{pq}{p^2+q^2} = 0$$

e

$$\lim_{q\to 0} \frac{pq}{p^2+q^2} = 0$$

Tuttavia queste sono solo due delle possibile successioni che si possono prendere. Prendiamo ad esempio sempre una successione che passa su una retta, ma che questa volta è nella forma $\vec{z}_n = (n, n)$*, ovvero la retta bisettrice del primo quadrante. In questo caso*

$$f(z_n) = \frac{nn}{n^2+n^2} = \frac{1}{2}$$

Il cui limite, per $n \to \infty$*, è ovviamente sempre* $1/2$*. Prendendo poi diverse rette si possono ottenere ancora risultati diversi. Questo mostra che, dato che i singolo limiti sono diversi, la funzione nel punto dato non presenta limite, al contrario di quanto poteva apparire ad una analisi affrettata.*

Ecco allora che, non esistendo questo limite, effettivamente nel punto $(0, 0)$ *la funzione* non possiede piano tangente*, pur essendo derivabile.*

In realtà esistono sistemi più semplici della verifica del limite dato per vedere se la funzione ammette nel punto (x_0, y_0) piano tangente. Ad esempio, il metodo più usato è il seguente:

Teorema 9.3.2 *Se la funzione* $f : A \in \mathbb{R}^n \to \mathbb{R}$ *possiede in un intorno del punto* (x_0, y_0) *derivate parziali prime continue, allora è differenziabile nel punto* (x_0, y_0)*.*

Questa è una condizione *sufficiente*, ma non *necessaria* – il che significa che, se una funzione soddisfa le condizioni date, sarà sicuramente differenziabile nel punto dato, e quindi ammetterà piano tangente, ma non è detto che se non le soddisfa sia obbligatoriamente non differenziabile.

Una funzione continua in A si dice che appartiene alla classe $C^0(A)$. Una funzione che che ammetta derivate prime continue in questo insieme appartiene invece alla classe $C^1(A)$. Andando avanti, una funzione che possieda derivate *seconde* continue apparterrà alla classe $C^2(A)$: $f \in C^2(A)$. Una funzione derivabile infinite volte in A (e quindi con tutte le derivate continue) apparterrà alla classe $C^\infty(A)$ ed è il miglior caso sperabile. La funzione esaminata nell'esempio precendente invece apparteneva a $C^0(\mathbb{R}^2)$, poichè non era vero che le sue derivate prime erano anche continue in tutto $\mathbb{R}^2$.

9.3.3 Gradiente e formula di Taylor al primo grado

Una entità molto utile nello studio delle funzioni a più variabili è il *gradiente*. Esso è un vettore che ha per componenti le derivate parziali della funzione in questione e si indica nel modo seguente:

$$\vec{\nabla} f(\vec{x}) = \frac{\partial f}{\partial x_1}(x,y)\vec{e}_1 + \frac{\partial f}{\partial x_2}(x,y)\vec{e}_2 + \ldots + \frac{\partial f}{\partial x_n}(x,y)\vec{e}_n = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x,y)\vec{e}_i$$

Quindi, il gradiente formalmente è una funzione vettoriale a valori vettoriali.

Già un suo utilizzo, senza volere, l'abbiamo introdotto; infatti abbiamo detto che dove si trova un massimo o minimo, tutte le derivate parziali prime si annullano, il che è come dire

$$\text{in } (x_0, y_0) \text{ è presente un max/min} \implies \vec{\nabla} f(x_0, y_0) = \vec{0}$$

Il gradiente possiede anche una interpretazione geometrica.

Inanzitutto, diciamo che le *curve di livello* di una funzione a due variabili sono quelle curve su cui $f(x, y) = k$. Ovviamente, se ci si trova in $\mathbb{R}^3$ si avranno superfici di livello, e così via.

Ora, per "scendere" o "salire" il minimo possibile (ovvero non farlo affatto) a partire da un certo punto $\vec{p} = (x_0, y_0)$, basterà seguire le curve di livello, in quanto, presa una qualunque direzione $\vec{n}$, se ci si muove sulla retta

$$\vec{x} = \vec{p} + t\vec{n}, \quad t \in \mathbb{R}$$

prendendo per $\vec{n}$ sempre la direzione della curva di livello, il valore della funzione rimarrà costante, e quindi la differenza sarà nulla.

Invece per trovare la direzione di massimo cambiamento, quale è la direzione da seguire?

Cerchiamo inanzitutto di scrivere la funzione in un altro modo. Se la funzione considerata è differenziabile, allora avevamo scritto che

$$\Delta = f(x_0 + h_x, y_0 + h_y) - f(x_0, y_0) - \frac{\partial f}{\partial x}(x_0, y_0)h_x - \frac{\partial f}{\partial y}(x_0, y_0)h_y$$

e

$$\frac{\Delta}{|\vec{h}|} \to 0$$

Ma questo significa che $\Delta = o(|\vec{h}|)$, e la definizione, ora che possediamo il concetto di gradiente, la possiamo riscrivere come:

$$f(\vec{x}_0 + \vec{h}) - f(\vec{x}_0) - \overrightarrow{\nabla} f(\vec{x}_0) \cdot \vec{h}$$

E, riassumendo il tutto in un'unica formula, otteniamo il polinomio di Taylor di primo grado di una funzione a più variabili:

$$f(\vec{x} + \vec{h}) - f(\vec{x}) = \overrightarrow{\nabla} f(\vec{x}) \cdot \vec{h} + o(|\vec{h}|)$$

Che coincide con la vecchia formula

$$f(x + h) - f(x) = f'(x)h + o(h)$$

Quindi la corretta trasposizione del concetto di derivata prima in più dimensioni è quello del gradiente.

A questo punto, applichiamo quello che abbiamo trovato: trascurando il termine infinitesimale $o(|\vec{h}|)$, possiamo scrivere che

$$f(\vec{x} + \vec{h}) - f(\vec{x}) \approx \overrightarrow{\nabla} f(\vec{x}) \cdot \vec{h}$$

E quindi, ovviamente, la direzione in cui il dislivello della funzione è maggiore è quello in cui il prodotto vettoriale è maggiore. Esso è massimo quando l'angolo tra i due vettori è nullo o di π: nel primo caso quindi $\vec{h}$ deve seguire la direzione del gradiente, nell'altro deve andare nel verso opposto. È facile trovare che nel primo caso si ha il massimo aumento della funzione, e nell'altro caso la massima diminuzione. Ecco trovata l'interpretazione geometrica del gradiente.

Considerando inoltre che la curva di livello seguirà ovviamente nella direzione in cui il dislivello è nullo, ovvero in cui $\overrightarrow{\nabla} f(\vec{x}) \cdot \vec{h} = 0$, è facile verificare che il gradiente è sempre perpendicolare alle curve di livello.

Queste conclusioni sono in accordo con il teorema di Fermat, infatti in un punto di massimo o minimo non esiste nessuna direzione privilegiata nella quale muoversi per avere un aumento o diminuzione massima; infatti in questi punti il gradiente non da alcune informazione: $\overrightarrow{\nabla} f = \vec{0}$.

9.3.4 Derivate seconde

Come già detto, se una funzione possiede n derivate parziali prime:

$$\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \cdots \frac{\partial f}{\partial x_n}$$

Ognuna di queste, in linea teorica, potrà possedere n derivate parziali seconde; una generica derivata parziale seconda, dove la f è stata derivata prima secondo x_k quindi secondo x_j viene indicata con la seguente notazione:

$$\frac{\partial^2 f}{\partial x_j \partial x_k} = \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_k}\right)$$

O anche con

$$D_{x_j x_k} f$$

O con

$$f_{x_j x_k}$$

In genere, è bene notare che

$$\frac{\partial^2 f}{\partial x_j \partial x_k} \neq \frac{\partial^2 f}{\partial x_k \partial x_j}$$

Tuttavia esiste un importante lemma che da le condizioni necessarie perchè invece esse siano uguali:

Lemma 9.3.3 (Lemma di Schwarz) *Se $f \in C^2(A)$, allora $f \in C^1(A)$, e le sue derivate parziale seconde f_{xy} e f_{yx} sono uguali.*

Ovvero, se le derivate parziali seconde di una funzione sono continue, allora lo sono anche le derivate parziali prime, ed inoltre si ottengono risultati uguali differenziando la funzione prima secondo una variabile x e poi secondo una variabile j, o facendo il contrario.

Così come il gradiente è il modo standard per rappresentare insieme le derivate parziali prime di una funzione, e rappresenta la derivata prima in più dimensioni, così per rappresentare le derivate parziali seconde di una funzione si utilizza la cosiddetta *matrice hessiana*:

$$H_f = \begin{bmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_2 \partial x_1} & \cdots & \dfrac{\partial^2 f}{\partial x_n \partial x_1} \\ \dfrac{\partial^2 f}{\partial x_1 \partial x_2} & \dfrac{\partial^2 f}{\partial x_2^2} & \cdots & \dfrac{\partial^2 f}{\partial x_n \partial x_2} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial x_1 \partial x_n} & \dfrac{\partial^2 f}{\partial x_2 \partial x_n} & \cdots & \dfrac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

Ovvero

$$H_f = (h_{ij})$$

e

$$h_{ij} = \frac{\partial^2 f}{\partial x_j \partial x_i}$$

Quindi, la matrice hessiana è una cosiddetta *matrice funzionale*, ovvero una matrice che ha per elementi funzioni. Quindi, si potrebbe dire che è una funzione che ha per variabile indipendente un vettore di $\mathbb{R}^n$ e come variabile dipendente una matrice $n \times n$.

Se le derivate parziali seconde sono continue, per il lemma di Schwarz la matrice hessiana è simmetrica, ovvero:

$$H_f^T = H_f$$

A questo punto possiamo scrivere, senza dimostrare questa volta, la formula di Taylor al second'ordine per le funzioni reali a più variabili:

$$f(\vec{x}_0 + \vec{h}) = f(\vec{x}_0) + \overrightarrow{\nabla} f(\vec{x}_0) \cdot \vec{h} + \frac{1}{2} H_f(\vec{x}_0) \times \vec{h} \cdot \vec{h} + o(|\vec{h}|^2)$$

Ovviamente nel moltiplicare la matrice hessiana per il vettore $\vec{h}$, si considera quest'ultimo come un vettore colonna. Avremo quindi $H_{n\times n} \times h_{n\times 1} = A_{n\times 1}$, e quindi ancora un vettore. A questo punto si moltiplica il vettore risultante per il vettore $\vec{h}$ ottenendo quindi uno scalare.

A questo punto possiamo fare alcune considerazioni. Ad esempio, in un punto in cui $\overrightarrow{\nabla} f(x_0) = 0$, ovvero in cui potrebbe esserci un massimo o minimo, la formula si riduce a

$$f(\vec{x}_0 + \vec{h}) - f(\vec{x}_0) = \frac{1}{2} H_f(\vec{x}_0) \times \vec{h} \cdot \vec{h} + o(|\vec{h}|^2)$$

Ora, se il termine contenente l'hessiana è sempre positivo, si ha, trascurando l'infinitesimo, $\forall \vec{h} \neq \vec{0}$, che

$$f(\vec{x}_0 + \vec{h}) - f(\vec{x}_0) > 0$$

e quindi in $\vec{x}_0$ c'è un minimo.

Ora, per un criterio rigoroso, prendiamo per partire il caso in cui $n = 2$ e le derivate parziali seconde continue, quindi la matrice hessiana simmetrica. Scriviamola allora in questa forma:

$$H_f(\vec{x}_0) = \begin{bmatrix} \alpha & \beta \\ \beta & \gamma \end{bmatrix}$$

Quindi, se poniamo $\vec{h} = (h_1, h_2)$:

$$H_f(\vec{x}_0) \times \vec{h} = \begin{bmatrix} \alpha & \beta \\ \beta & \gamma \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} = \begin{bmatrix} \alpha h_1 + \beta h_2 \\ \beta h_1 + \gamma h_2 \end{bmatrix}$$

E quindi

$$\left[H_f(\vec{x}_0) \times \vec{h} \right] \cdot \vec{h} \ = \ (\alpha h_1 + \beta h_2, \beta h_1 + \gamma h_2) \cdot (h_1, h_2)$$

Che, semplificando, diventa, molto semplicemente:

$$\alpha h_1^2 + 2\beta h_1 h_2 + \gamma h_2^2$$

Ora, dobbiamo trovare il modo di esprimere il fatto che questo valore deve essere maggiore di zero, ovvero dobbiamo risolvere la disequazione

$$\alpha h_1^2 + 2\beta h_1 h_2 + \gamma h_2^2 > 0$$

Per tutti gli $\vec{h} \neq \vec{0}$. Prendiamo prima l'insieme di casi, molto semplice, in cui $\vec{h} = (h_1, 0)$, ovvero $h_2 = 0$: la formula si riduce a

$$\alpha h_1^2 > 0$$

Ovvero

$$\alpha > 0$$

Adesso consideriamo invece il caso in cui $h_2 \neq 0$; quindi possiamo dividere entrambi i membri per h_2^2, che quindi è sempre maggiore di zero, e otteniamo:

$$\frac{\alpha h_1^2 + 2\beta h_1 h_2 + \gamma h_2^2}{h_2^2} > 0$$

Ovvero

$$\alpha \left(\frac{h_1}{h_2}\right)^2 + 2\beta \frac{h_1}{h_2} + \gamma > 0$$

Poniamo ora

$$t = \frac{h_1}{h_2}$$

Ed abbiamo la disequazione in una variabile

$$\alpha t^2 + 2\beta t + \gamma > 0$$

Ora, questa disequazione è vera per qualunque t solo se il suo determinante è minore di zero, ovvero se

$$\beta^2 - \alpha\gamma < 0$$

o anche

$$\alpha\gamma - \beta^2 > 0$$

Ma il membro di sinistra è il determinante della nostra matrice hessiana, e quindi, riassumendo tutte le condizioni trovate finora:

$$\begin{cases} \alpha > 0 \\ \det H_f(\vec{x}_0) > 0 \end{cases}$$

E queste sono le condizioni per avere un minimo relativo. In modo analogo si trovano le condizioni per avere un massimo relativo, ovvero:

$$\begin{cases} \alpha < 0 \\ \det H_f(\vec{x}_0) > 0 \end{cases}$$

Nel caso in cui invece il determinante della matrice hessiana sia minore di zero, significa che non si presenta nè massimo nè minimo: per convincersene basta rileggere la dimostrazione fatta.

Se prendiamo un caso più complesso, ad esempio con $n = 3$, la disequazione finale da risolvere risulta più complessa: ad esempio se indichiamo la matrice hessiana con

$$\begin{bmatrix} \alpha & \beta & \gamma \\ \beta & \omega & \eta \\ \gamma & \eta & \theta \end{bmatrix}$$

Allora essa sarà

$$\alpha h_1^2 + \omega h_2^2 + \theta h_3^2 + 2\beta h_1 h_2 + 2\gamma h_1 h_3 + 2\eta h_2 h_3 > 0$$

Ora, nel caso in cui $h_3 = 0$, ci troviamo nel caso già studiato, da cui ricaviamo

$$\begin{cases} \alpha > 0 \\ \det \begin{bmatrix} \alpha & \beta \\ \beta & \omega \end{bmatrix} > 0 \end{cases}$$

Ma per il caso generico la soluzione è alquanto più complessa da determinare – tuttavia di per sè non è difficile da esprimere, ed è la seguente:

$$\begin{cases} \alpha > 0 \\ \det \begin{bmatrix} \alpha & \beta \\ \beta & \omega \end{bmatrix} > 0 \\ \det \begin{bmatrix} \alpha & \beta & \gamma \\ \beta & \omega & \eta \\ \gamma & \eta & \theta \end{bmatrix} > 0 \end{cases}$$

Per dare la soluzione generale occorre prima introdurre la seguente notazione: con Δ_n si intende il determinante della matrice ottenuta prendendo le prime n righe ed n colonne della matrice hessiana. Detto ciò, si può dimostrare che:

Criterio 9.3.4 *É vero che*

1.
$$\forall \vec{h} \neq \vec{0} : H_f(\vec{x}_0) \times \vec{h} \cdot \vec{h} > 0 \iff \Delta_1 > 0, \Delta_2 > 0, \ldots, \Delta_n > 0$$

2.

$$\forall \vec{h} \neq \vec{0} : H_f(\vec{x}_0) \times \vec{h} \cdot \vec{h} < 0 \iff \Delta_1 < 0, \Delta_2 > 0, \Delta_3 < 0, \dots$$

Ovvero che

$$\Delta_{2n+1} < 0$$

e

$$\Delta_{2n} > 0$$

Se la sequenza dei segni non appartiene nè all'una nè all'altra configurazione, allora non si ha nè massimo nè minimo; se invece almeno uno dei Δ è nullo, non si può dire niente a riguardo.

Esempio 9.3.4 *Trovare massimi e minimi relativi della funzione*

$$f : \mathbb{R}^2 \to \mathbb{R};\ f(x,y) = x^2 + y^4$$

Possiamo inanzitutto osservare che, per tutte le $\vec{x} \neq \vec{0}$, si ha che $f(\vec{x}) > 0$, mentre $f(\vec{0}) = 0$, quindi sicuramente $(0,0)$ è un punto non solo di minimo ma addirittura di minimo assoluto.

Utilizziamo ora il teorema di Fermat per trovare i punti critici della funzione:

$$\overrightarrow{\nabla} f(x,y) = (2x, 4y^3)$$

Quindi ci troviamo a risolvere il sistema:

$$\begin{cases} 2x = 0 \\ 4y^3 = 0 \end{cases}$$

Da cui ricaviamo la soluzione che già conosciamo $(x,y) = (0,0)$. Vediamo ora lo studio della matrice hessiana che informazioni ci da:

$$\det H_f(0,0) = \det \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} = 0$$

E quindi noi non possiamo dire nulla riguardo la funzione nel punto $(0,0)$, in base alla matrice hessiana. Tuttavia, sappiamo, *che esiste lì un minimo. Come tutti i criteri quello della matrice hessiana è utile solo se si incontrano le ipotesi date.*

Esempio 9.3.5 *Calcolare minimi e massimi della funzione*

$$f : \mathbb{R}^2 \to \mathbb{R};\ f(x,y) = e^{x^2 - xy + 2y^2}$$

Come sempre, calcoliamo il gradiente della funzione:

$$\overrightarrow{\nabla} f(x,y) = \left[(2x - y)e^{x^2 - xy + 2y^2}, (4y - x)e^{x^2 - xy + 2y^2} \right]$$

Dato che le esponenziali sono sempre positive, il gradiente si annullerà semplicemente sotto queste condizioni:

$$\begin{cases} 2x - y = 0 \\ -x + 4y = 0 \end{cases} \quad \begin{cases} y = 2x \\ 8x - x = 0 \end{cases}$$

La cui unica soluzione è $(x,y) = (0,0)$. Calcolando le derivate seconde si ottiene:

$$\frac{\partial f}{\partial^2 x} = (4x^2 - 4xy + y^2 + 2)e^{x^2 - xy + 2y^2}$$

$$\frac{\partial f}{\partial^2 y} = (16y^2 - 8xy + x^2 + 16)e^{x^2 - xy + 2y^2}$$

$$\frac{\partial f}{\partial x \partial y} = [-1 + (4y - x)(2x - y)]e^{x^2 - xy + 2y^2}$$

Calcolate nel punto zero portano al seguente determinante della matrice hessiana:

$$\det \begin{bmatrix} 2 & -1 \\ -1 & 4 \end{bmatrix} = 8 + 1 > 0$$

Dal fatto poi che $\alpha > 0$, si ottiene che in $(0,0)$ è presente un minimo. Inoltre $f(0,0) = 1$, e sappiamo che è una funzione continua definita su un insieme, e quindi ha per codominio un intervallo. Questo intervallo avrà come minimo quindi 1. Inoltre è facile verificare che la funzione tende ad infinito per x o y che tendono ad infinito, e quindi:

$$f(\mathbb{R}) = [1; +\infty[$$

In particolare l'uso dello studio di massimi e minimi a più variabili può portare ad un interessante risultato spesso utilizzato in varie scienze, tra cui la fisica, ed è il *metodo dei minimi quadrati.*

Esempio 9.3.6 *Data un insieme di coppie*

$$(p_n, q_n)$$

Di valori reali distinti, con $n \geq 2$, suppongo che esista tra loro una relazione lineare, sebbene questi valori si possano discostare di poco dalla ideale retta su cui tutti dovrebbero stare. Trovare quindi i coefficienti della retta

$$y = ax + b$$

tale per cui sia minima la distanza dell'insieme dei punti dalla retta.

Questo problema di ottimizzazione porta in gioco la somma delle distanza dei punti dalla retta:

$$\sum_{i=1}^{n} |(p_i, q_i) - (p_i, ap_i + b)|$$

E il nostro scopo sarà ovviamente quello di rendere tale quantità minima. Ora, la distanza tra un punto generico (p_i, q_i) e la retta $y = ax + b$ è dato da:

$$g_i(a, b) = \frac{|ap_i + b - 1_i|}{\sqrt{a^2 + b^2}}$$

E quindi noi dobbiamo trovare il minimo di

$$\sum |ap_i + b - q_i|$$

Ovviamente, dove tali quantità saranno minime, lo saranno anche i loro quadrati, che però risultano di più comodo utilizzo:

$$F(a, b) = \sum_{i=1}^{n} [ap_i + b - q_i]^2$$

La funzione F così introdotta viene chiamata scarto quadratico medio. *Ovviamente il dominio di F è l'intero $\mathbb{R}^{\nvDash}$, quindi ricerchiamo prima i punti critici.*

$$\frac{\partial F}{\partial a} = \sum_{i=1}^{n} 2p_i(ap_i + b - q_i)$$

$$\frac{\partial F}{\partial b} = \sum_{i=1}^{n} 2(ap_i + b - q_i)$$

Ora eguagliamo le due derivate prime a zero (vengono omessi gli indici delle sommatorie perchè sempre uguali a quelli precedenti):

$$\begin{cases} a\sum p_i^2 + b\sum p_i = \sum p_i q_i \\ a\sum p_i + nb = \sum q_i \end{cases}$$

Ora abbiamo quindi un sistema lineare, la cui matrice dei coefficienti è:

$$A = \begin{bmatrix} \sum p_i^2 & \sum p_i \\ \sum p_i & n \end{bmatrix}$$

Noi sappiamo che il nostro sistema ammette soluzioni solo se $\det A \neq 0$*:*

$$\det A = n\sum p_i^2 - (\sum p_i)^2 > 0$$

Questo è sempre vero poichè è sempre vero che, se presi $p_1, p_2, \ldots, p_n$, con $p_k \neq p_{k+1}$, allora

$$(\sum p_i)^2 < n\sum p_i^2$$

Non faremo la dimostrazione per n qualunque, tuttavia ce ne si può convincere utilizzando, ad esempio, $n = 2$, per il quale si ha che:

$$(p_1 + p_2)^2 < 2(p_1^2 + p_2^2)$$

$$p_1^2 + 1p_1p_2 + p_2^2 < 2p_1^2 + 2p_2^2$$

$$-p_1^2 - p_2^2 + 2p_1p_2 < 0$$

$$(p_1 - p_2)^2 > 0$$

E quindi l'uguaglianza è vera.

Tornando alla nostra funzione, dato che il determinante è sempre maggiore di zero, so per certo che il sistema ha una ed una sola soluzione, e quindi F uno ed un solo punto critico. Utilizzando l'hessiana:

$$\frac{\partial F}{\partial^2 a} = \sum p_i^2$$
$$\frac{\partial F}{\partial a \partial b} = \frac{\partial F}{\partial b \partial a} = \sum p_i$$
$$\frac{\partial F}{\partial^2 b} = n$$

E quindi la matrice hessiana vale:

$$H_F = \begin{bmatrix} \sum p_i^2 & \sum p_i \\ \sum p_i & n \end{bmatrix} = A$$

Quindi, per le stesse ragioni di prima, il suo determinante sarà sempre maggiore di zero. Quindi, dato che il determinante della matrice hessiana è maggiore di zero ed $\alpha > 0$, abbiamo nel punto preso in considerazione un minimo. Questo punto sarà quindi la soluzione del sistema dato prima, poco interessante da calcolare, ma intanto abbiamo dimostrato l'esistenza della cosiddetta retta di regressione *per qualunque insieme di punti a due a due distinti.*

9.4 Cenni di topologia in $\mathbb{R}^n$

Valgono ancora i principali teoremi sulle funzioni continue dati per $f : \mathbb{R} \to \mathbb{R}$? La risposta è positiva, ma con le dovute modifiche che permettono di generalizzare il loro significato. Per fare questo dovremo introdurre alcune nozioni di topologia in $\mathbb{R}^n$.

Il primo concetto è quello di insieme *limitato*. Un insieme $a \subset \mathbb{R}^n$ si dice limitato se e solo se $\exists r > 0 : \forall \vec{x} \in A : |\vec{x}| \leq r$, ovvero se l'intero insieme è contenuto all'interno di una sfera n-dimensionale di raggio r. Se trasportiamo questa definizione in $\mathbb{R}$ otteniamo ancora la definizione di insieme limitato trattato in precedenza.

Dire invece che $A \subset \mathbb{R}$ è *chiuso* significa esiste un numero *finito* di funzioni $\Phi_1, \Phi_2, \ldots, \Phi_n$ da $\mathbb{R}^n$ ad $\mathbb{R}$ e qui continue tali per cui si possa riscrivere l'insieme A nella seguente forma:

$$A = \{\vec{x} \in \mathbb{R}^n \mid \Phi_1(\vec{x}) \geq 0, \Phi_2(\vec{x}) \geq 0, \ldots, \Phi_n(\vec{x}) \geq 0\}$$

Esempio 9.4.1 *Determinare se l'insieme*

$$A = \left\{(x, y) \in \mathbb{R}^2 \mid y \geq x^2, x + y \leq 3\right\}$$

Se prendiamo le funzioni:

$$\Phi_1(x, y) = y - x^2$$
$$\Phi_2(x, y) = 3 - x - y$$

Possiamo effettivamente riscrivere l'insieme come:

$$A = \left\{(x, y) \in \mathbb{R}^2 \mid \Phi_1(x, y) \geq 0, \Phi_2(x, y) \geq 0\right\}$$

Quindi l'insieme dato è chiuso, ed è inoltre anche limitato.

Esempio 9.4.2 *Determinare se l'insieme*

$$A = \left\{(x, y, z) \in \mathbb{R}^3 \mid y \geq 0, x^2 + y^2 = z^2\right\}$$

In questo caso prendiamo le funzioni:

$$\Phi_1(x, y, z) = y$$
$$\Phi_2(x, y, z) = x^2 + y^2 - z^2$$
$$\Phi_2(x, y, z) = -x^2 - y^2 + z^2$$

Da notare che l'unione delle condizioni imposte dalle due ultime funzioni forma la seconda condizione dell'insieme.

A questo punto possiamo definire di nuovo il teorema di Weierstrass:

Teorema 9.4.1 (Teorema di Weierstrass in $\mathbb{R}^n$) *Data una $f : A \subset \mathbb{R}^n$, se A è chiuso e limitato ed f è continua su A, allora la funzione ammette massimo e minimo assoluto.*

Per definire invece gli altri due teoremi fondamentali, quello degli zeri e di Bolzano, occorre trovare la nozione corrispondente all'intervallo di $\mathbb{R}$. Quello che cerchiamo prende il nome di insieme convesso. Un insieme A si dice *connesso* se e solo se, per qualsiasi coppia $\vec{a} \neq \vec{b}$ di punti appartenenti ad A, esiste una poligonale interamente contenuta in A che li unisce. Con poligonale si intende una serie di segmenti uniti a due a due ad un estremo. Si dice inoltre *insieme convesso* un insieme che, presi $\vec{a} \neq \vec{b}$ contenuti nell'insieme, il segmento che li congiunge è anch'esso contenuto nell'insieme.

Ovviamente qualunque insieme convesso è anche connesso, ma non viceversa. L'unione di due insiemi connessi è ancora un insieme connesso, se i due insiemi hanno almeno un punto in comune, ma l'unione di due insiemi convessi non è detto che sia convesso (anche se senz'altro connesso).

Possiamo ora enunciare anche gli ultimi due teoremi:

Teorema 9.4.2 (Teorema degli zeri in $\mathbb{R}^n$) *Data*

$$f : A \subset \mathbb{R}^n \to \mathbb{R}$$

con A insieme connesso ed f continua, si ha che, presa una coppia di punti di A $\vec{a} \neq \vec{b}$, con $f(\vec{a})f(\vec{b}) < 0$, allora su qualunque poligonale interna ad A che congiunge questi due punti la funzione si annulla in almeno un punto.

Da notare che questa volta non abbiamo garantita l'esistenza solo di uno zero, ma di almeno tanti quante sono le poligonali costruibili, in che spesso significa infinite!

Per quanto riguarda l'ultimo:

Teorema 9.4.3 (Teorema di Bolzano (o del valore intermedio) in $\mathbb{R}^n$) *Data*

$$f : A \subset \mathbb{R}^n \to \mathbb{R}$$

con A connesso ed f continua, si ha che o la funzione data è costante, oppure $f(A)$ è un intervallo.

Esempio 9.4.3 *Dato*

$$A = \left\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 4\right\}$$

E

$$f : A \to \mathbb{R} \ : \ x^2 - xy + y^2$$

Determinare f(A)

L'insieme A è ovviamente il cerchio di centro l'origine e raggio 2. Questo insieme è chiuso (basta prendere $\Phi(x, y) = x^2 - y^2 - 4$), limitato (ovviamente $r = 2$; $\forall a \in A \ : \ |a| \leq 2$) e connesso (anzi, di più, convesso, e ciò si può dimostrare con semplici proprietà geometriche del cerchio). Inoltre, la f è continua su tutto $\mathbb{R}^2$, quindi anche su $A \subset \mathbb{R}^2$. Questo permette di dire, per il teorema di Weierstrass, che essa possiede massimo e minimo assoluti. Il teorema di Bolzano invece ci dice che $f(A)$ è un intervallo, e quindi il problema dato si riconduce a trovare

$$\left[\min_A f; \max_A f\right]$$

Che è un problema che sappiamo risolvere.

Prima di tutto semplifichiamo la funzione in:

$$f(x, y) = (x - y)^2 + xy$$

E calcoliamo la derivata prima:

$$f_x(x, y) = 2(x - y) + y$$
$$f_y(x, y) = -2(x - y) + x$$

Quindi il sistema:

$$\begin{cases} f_x(x, y) = 0 \\ f_y(x, y) = 0 \end{cases}$$

ha come unica soluzione $(x, y) = (0, 0)$. Qui la Hessiana vale:

$$\det \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} = 5 > 0$$

E quindi, dato che $\alpha > 0$, si ha un minimo.

A questo punto ci troviamo in una strana situazione: abbiamo apparentemente trovato il minimo della nostra funzione, ma... dove si trova il massimo? Ovviamente non può essere lo stesso punto di prima, in quanto la funzione non è costante.

L'errore che abbiamo commesso è che, utilizzando il teorema di Fermat, bisogna tenere conto di una condizione che al tempo non considerammo perchè mancava di definizione, ma che adesso è stata data, ovvero che A deve essere un insieme aperto. *Se prendiamo invece il nostro insieme A, esso risulta chiuso, come dicevamo prima!*

A quanto pare siamo in un vicolo cieco per quanto riguarda la ricerca di un codominio di una funzione su un insieme: se questo insieme è aperto non possiamo utilizzare il teorema di Weierstrass, se invece è chiuso non possiamo utilizzare quello di Fermat.

In realà la soluzione esiste. Se prendiamo infatti l'insieme in questione e togliamo da esso la sua frontiera, esso diventa:

$$A' = \left\{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 < 4\right\}$$

(notare il cambiamento da $\leq$ a $<$).

A questo punto si possono cercare massimi e minimi relativi su questo insieme con il teorema di Fermat, ed in effetti abbiamo trovato la soluzione: in questo insieme la funzione ammette minimo, ma non massimo evidentemente. Quindi, la ricerca a questo punto si sposta proprio sulla frontiera di A, ovvero:

$$\partial A = \left\{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 = 4\right\}$$

Se riusciamo a trovare il modo di compiere una ricerca di massimo e minimo su insiemi di questo tipo, basterà a questo punto unire i due insiemi di soluzioni per ottenere il minimo assoluto e il massimo assoluto. La ricerca di massimi e minimi su questi particolari insiemi, chiamati vincoli*, viene detta "ricerca di* massimi e minimi vincolati*".*

9.5 Massimi e minimi vincolati

Cominciamo inanzitutto con qualche considerazione. Prendiamo una generica funzione

$$f : A \subseteq \mathbb{R}^n \to \mathbb{R}$$

ed un'altra funzione, che chiameremo *vincolo*:

$$g : A \subseteq \mathbb{R}^n \to \mathbb{R}$$

E diciamo che cerchiamo il massimo e minimo di f sull'insieme definito dai punti dove $g(\vec{x}) = 0$. Quindi, cerchiamo i punti $\vec{x}_0$ per i quali vale che:

$$\exists r \in \mathbb{R} \ : \ \forall \vec{x} \in \{\vec{v} \mid g(\vec{v}) = 0, |\vec{v} - \vec{x}| < r, \vec{v} \neq \vec{x}_0\} \ : \ f(\vec{x}) < (\text{ oppure } >) f(\vec{x}_0)$$

Un primo metodo di soluzione potrebbe essere il seguente: si esplicita l'equazione di $g(\vec{x}) = 0$ per uno dei componenti, e si compie una sostituzione sulla funzione originale. A questo punto è facile convincersi che il problema si riconduce a trovare un massimo o minimo *libero* sulla funzione così descritta, che ha una variabile in meno. Questo metodo è molto comodo quando il vincolo risulta semplice, e, soprattutto, la sua forma esplicitata risulta altrettanto semplice.

Esempio 9.5.1 *Calcolare massimi e minimi relativi della funzione*

$$f : \mathbb{R}^2 \to \mathbb{R} \ : \ x^2 - \ln y$$

Sul vincolo

$$g(x,y) = y - e^x \ ; \ g(x,y) = 0$$

Esplicitando il vincolo per, ad esempio, la y, otteniamo:

$$y = e^x$$

E, sostituendo nella f:

$$f(x) = x^2 - \ln(e^x) = x^2 - x$$

Si ottiene facilmente che questa funzione ha minimo in $x = 1/2$ e nessun massimo. Quindi, otteniamo che la corrispondente y vale $e^{1/2}$. La soluzione cercata è quindi: esiste un minimo nel punto

$$\vec{x}_0 = \left(\frac{1}{2}, \sqrt{e}\right)$$

Tuttavia questo metodo spesso non è applicabile in presenza di funzioni più complesse. Possiamo comunque compiere alcune osservazioni.

Prendiamo un qualunque punto $\vec{x}_0$ sull'insieme definito dal vincolo, e disegnamo la curva di livello della nostra funzione f passante per $\vec{x}_0$. Diciamo che su questa curva di livello la funzione valga k. Per semplicità, assumiamo di star lavorando con una $f : \mathbb{R}^2 \to \mathbb{R}$, anche se i ragionamenti qui fatti non cambiano a seconda del numero di dimensioni considerate.

Disegnamo ora oltre la curva di livello anche la curva $g(x,y) = 0$, che rappresenta quindi l'insieme su cui stiamo cercando massimi e minimi. Le due curve potranno quindi essere tra loro tangenti oppure secanti in $\vec{x}_0$. Per esprimere meglio la condizione di tangenza, diciamo che le due curve avranno tangente uguale nel punto $\vec{x}_0$ oppure, che è la stessa cosa, i rispettivi gradienti linearmente dipendenti (in quanto i gradienti sono sempre perpendicolari alle curve di livello, e quindi alle tangenti ad esse).

Se le due rette sono secanti, non si potrà ovviamente avere mai un minimo o massimo. Prendiamo infatti un intorno del punto $\vec{x}_0$. Ora, nella sezione dell'intorno da un parte della curva di livello avremo valori di $f(\vec{x}) > k$, dall'altra parte invece

$f(\vec{x}) < k$. Ma dato che il vincolo passa da entrambe le parti, questo significa che il valore della funzione, sempre seguendo il vincolo, passa da minore di $f(\vec{x}_0)$, a maggiore di esso. Quindi, non può esserci massimo o minimo relativo.

Se invece la curva di livello è tangente, si avrà sicuramente, per lo stesso ragionamento, che la funzione ha valori tutti maggiori o minori di k, e quindi avremo un massimo o minimo relativo.

Ecco quindi che abbiamo trovato la soluzione che cercavamo. I punti di massimo o minimo relativo su un vincolo sono quei punti (appartenenti al vincolo) per i quali si abbia che i gradienti della funzione e della funzione di vincolo sono proporzionali (per due vettori, la dipendenza lineare è in realtà una semplice questione di multipli).

Esponendo in maniera formale ciò che abbiamo detto, otteniamo che, data una funzione

$$f|_C : \mathbb{R}^n \to \mathbb{R}$$

Che si legge "f ristretto a C", dove C è l'insieme definito dal vincolo

$$g : \mathbb{R}^n \to \mathbb{R} \; ; \; g(\vec{x}) = 0$$

I punti di massimo e minimo relativo si trovano tra le soluzioni di:

$$\begin{cases} g(\vec{x}) = 0 \\ \\ \overrightarrow{\nabla} f(\vec{x}) = \lambda \overrightarrow{\nabla} g(\vec{x}) \end{cases}$$

O anche, scritto in forma estesa:

$$\begin{cases} g(\vec{x}) = 0 \\ f_{x_1}(x_1) = \lambda g_{x_1}(x_1) \\ f_{x_2}(x_2) = \lambda g_{x_2}(x_2) \\ \dots \\ f_{x_n}(x_n) = \lambda g_{x_n}(x_n) \end{cases}$$

Questo sistema di risoluzione del problema viene chiamato *metodo dei moltiplicatori di Lagrange*, poichè la variabile λ utilizzata ha nome, per l'appunto, di *moltiplicatore di Lagrange*.

Esempio 9.5.2 *Sia dato un punto materiale di massa m libero di muoversi solo su una circonferenza di raggio R. Questo punto è soggetto alla forza gravitazionale ed alla forza elastica di una molla di costante k agganciata ad un punto della circonferenza. Si trovino le posizioni di equilibrio stabile e instabile del sistema.*

Una posizione di equilibrio stabile è una posizione in cui la funzione che rappresenta l'energia totale in funzione della posizione risulta avere un minimo, mentre una posizione di equilibrio stabile si verifica in presenza di un massimo della stessa funzione.

Si sa anche che l'energia gravitazione di un corpo posto a quota h dallo zero che si prende, è $E = mgh$. La forza elastica è invece pari a $\frac{k}{2}d^2$, dove k è la costante elastica e d la distanza dalla posizione di equilibrio.

Con queste informazioni si può ora porre il problema. Fissiamo inanzitutto un sistema di coordinate cartesiane per porre gli elementi dati: Ad esempio si potrebbe prendere per origine il centro del cerchio. In questo caso abbiamo come energia gravitazionale:

$$E_g = mgh = mgy$$

L'elastico invece sarà lo agganciamo al punto $(R, 0)$ e quindi l'energia elastica sarà:

$$E_e = \frac{k}{2}d^2 = \frac{k}{2}[(x - R)^2 + y^2]$$

E l'energia totale sarà la somma di quelle singole. A questo punto abbiamo anche che:

$$x^2 + y^2 - R^2 = 0$$

(il punto non é libero), e dobbiamo trovare su questo vincolo i massimi e minimi della funzione data. Il problema si è quindi ricondotto ad una ricerca di massimo e minimo vincolato. Calcoliamo quindi i gradienti della funzione $E(x, y)$ e $g(x, y)$:

$$E_x = k(x - R)$$

e

$$E_y = mg + ky$$

Mentre, per quanto riguarda $g(x, y)$:

$$g_x = 2x$$

$$g_y = 2y$$

E quindi il sistema che dobbiamo risolvere è:

$$\begin{cases} x^2 + y^2 = R^2 \\ k(x - R) = 2\lambda x \\ mg - ky = 2\lambda y \end{cases}$$

Dalla seconda equazione ricaviamo la x:

$$x = \frac{kR}{k - 2\lambda}$$

Mentre dalla terza otteniamo y:

$$y = \frac{mg}{k + 2\lambda}$$

E infine, sostituendo nella prima, ricaviamo anche λ, e da qui le due x e y:

$$x = \pm \frac{kR}{\sqrt{m^2 g^2 + k^2 R^2}}$$

$$y =$$

***TODO**: Finire questo esempio!*

Quindi, ora abbiamo tutti gli strumenti necessari per determinare codomini di funzioni continue definite su insiemi chiusi.

9.6 Integrali doppi e tripli

9.6.1 Definizione di integrale doppio

Quando parlammo di integrazione in una variabile, ci trovammo a definire l'integrale come limite di somme di Cauchy:

$$\int_a^b f(x)\,\mathrm{d}x = \lim_{n\to\infty} \sum_{k=1}^{n} \frac{b-a}{n} f(\xi_k)$$

Dove

$$\xi_k \in \left[a + (k-1)\frac{b-a}{n}; a + k\frac{b-a}{n}\right]$$

E in questo modo riuscivamo ad esprimere l'area sottesa dalla funzione f nell'intervallo $[a; b]$. Quando ci troviamo ad utilizzare, ad esempio, funzioni a due variabili, la naturale estensione del concetto di integrale ci porta alla ricerca di un *volume* sotteso dal grafico di una funzione, che è una superficie.

Prendiamo come corrispondente dell'intervallo $[a; b]$, un rettangolo $[a; b] \times [c; d]$, e dividiamolo in n^2 intervallini:

$$I_{ij} = \left[a + (i-1)\frac{b-a}{n}; a + i\frac{b-a}{n}\right] \times \left[c + (j-1)\frac{d-c}{n}; c + j\frac{d-c}{n}\right]$$

Il che è come costruire una "scacchiera" con n caselle orizzontali ed n verticali. A questo punto prendiamo una serie di $\xi_{ij} \in I_{ij}$, e consideriamo la somma:

$$S(n) = \sum_{i,j=1}^{n} |I_{ij}| f(\xi_{ij})$$

Dipendente da n. In questo caso la somma trovata rappresenterà un'approssimazione del volume sotteso dalla superficie di equazione $z = f(x, y)$. Ora, all'aumentare di n, la nostra "scacchiera" si infittirà, e quindi il volume diventerà man mano più preciso. Definiamo allora la nozione di *integrale doppio su* $[a; b] \times [c; d]$ come il seguente limite:

$$\iint_{[a;b]\times[c;d]} f(x, y)\,\mathrm{d}x\,\mathrm{d}y = \lim_{n\to\infty} S(n)$$

Così come per le funzioni ad una variabile, anche qui vale che:

Teorema 9.6.1 *Sia $f : [a; b] \times [c; d] \to \mathbb{R}$ una funzione continua; in questo caso, il limite*

$$\lim_{n\to\infty} S(n)$$

Esiste finito.

Così facendo però ci siamo ristretti ad un caso molto specifico, ovvero quello in cui il cosiddetto *dominio di integrazione* è un rettangolo. È tuttavia facile definire cosa si intende per integrale doppio su un qualunque insieme limitato.

Dato l'insieme $\Omega \subset \mathbb{R}^2$ sul quale vogliamo integrare, prendiamo un qualunque rettangolo $[a; b] \times [c; d]$ che contenta Ω (e questo possiamo farlo perchè Ω è limitato) e la funzione seguente:

$$f^*(x, y) = \begin{cases} f(x, y) \text{ se } (x, y) \in \Omega \\ 0 \text{ se } (x, y) \notin \Omega \end{cases}$$

E definiamo così l'integrale doppio:

$$\iint_\Omega f(x, y)\,\mathrm{d}x\,\mathrm{d}y = \iint_{[a;b]\times[c;d]} f^*(x, y)\,\mathrm{d}x\,\mathrm{d}y$$

Tuttavia non è detto che una funzione continua sia integrabile su un qualsiasi dominio limitato. Se prendiamo ad esempio un rettangolo $[0; 1] \times [0; 1]$ costituito solo da punti però di $\mathbb{Q}^2$, è facile convincersi che una qualunque funzione non può essere integrata su questo intervallo. Ci conviene quindi considerare un insieme più ristretto dei domini utilizzabili.

9.6.2 Insiemi x-semplici ed y-semplici, domini regolari

Partiamo con:

Definizione 9.6.1 *Un insieme $A \subset \mathbb{R}^2$ si dice* y-semplice *se si può scrivere nella seguente forma:*

$$A = \left\{(x,y) \in \mathbb{R}^2 \ : \ x \in [a;b],\ g_1(x) \leq y \leq g_2(x)\right\}$$

con $g_1, g_2 : [a;b] \to \mathbb{R}$ funzioni continue, e $\forall x \in [a;b] : g_1(x) \leq g_2(x)$, ovviamente.

Esaminiamo la definizione data. In pratica l'insieme E risulta definito su un intervallo delle x, e, se tagliato su ognuna di queste x_0, abbiamo un segmento verticale. L'insieme così costruito è sicuramente limitato.

Insieme a questa definizione, ne esiste una speculare, quella degli insiemi *x-semplici*:

Definizione 9.6.2 *Un insieme $A \subset \mathbb{R}^2$ si dice* x-semplice *se si può scrivere nella seguente forma:*

$$A = \left\{(x,y) \in \mathbb{R}^2 \ : \ y \in [a;b],\ g_1(y) \leq x \leq g_2(y)\right\}$$

con $g_1, g_2 : [a;b] \to \mathbb{R}$ funzioni continue, e $\forall y \in [a;b] : g_1(y) \leq g_2(y)$, ovviamente.

Questo insieme non è altro che il simmetrico rispetto alla bisettrice del primo quadrante di un insieme y-semplice. Anche in questo caso, quindi, il modo più intuitivo per riconoscere un insieme di questo genere è vedere se una sezione dell'insieme in questione fatta rispetto alla y è o meno un segmento, e tali segmenti cambiano con continuità.

Infine:

Definizione 9.6.3 *Si dice* insieme regolare *l'unione di un numero finito di insiemi x o y-semplici senza punti interni in comune a due a due.*

***TODO**: Devono effettivamente avere punti a due a due in comune?*

Esempio 9.6.1 *L'insieme*

$$A = \left\{(x,y) \in \mathbb{R}^2 \ : \ 0 \leq x \leq 1 \ ; \ x(1-x) \leq y \leq x(4-x)\right\}$$

È Un insieme $y-$semplice, per definizione, ponendo $g_1(x) = x(1-x)$ e $g_2(x) = x(4-x)$.

Esempio 9.6.2 *L'insieme*

$$A = \left\{(x,y) \in \mathbb{R}^2 \ : \ x \in [0;2] \ ; \ 0 \leq y \leq \frac{x}{2}\right\}$$

È sicuramente y-semplice, ma è anche y-semplice, infatti:

$$A = \left\{(x,y) \in \mathbb{R}^2 \ : \ y \in [0;1] \ ; \ 2y \leq x \leq 2\right\}$$

Esempio 9.6.3 *Si consideri una corona circolare di raggio interno r e raggio esterno R:*

$$A = \left\{(x,y) \in \mathbb{R}^2 \ : \ r^2 \leq x^2 + y^2 \leq R^2\right\}$$

Disegnandola, ci si può convincere che questo non è nè un insieme x-semplice, nè un insieme y-semplice. È tuttavia sempre regolare, infatti è visualizzabile come l'unione di più insiemi x-semplici oppure y-semplici. Ad esempio:

$$A_1 = \left\{(x,y) \in \mathbb{R}^2 \ : \ y \in [-R;-r] \ ; \ x \in \left[-\sqrt{R^2-y^2}; \sqrt{R^2-y^2}\right]\right\}$$

$$A_2 = \left\{(x,y) \in \mathbb{R}^2 \ : \ y \in [-r;r] \ ; \ x \in \left[-\sqrt{R^2-y^2}; -\sqrt{r^2-y^2}\right]\right\}$$

$$A_3 = \left\{(x,y) \in \mathbb{R}^2 \ : \ y \in [-r;r] \ ; \ x \in \left[\sqrt{r^2-y^2}; \sqrt{R^2-y^2}\right]\right\}$$

$$A_4 = \left\{(x,y) \in \mathbb{R}^2 \ : \ y \in [r;R] \ ; \ x \in \left[-\sqrt{R^2-y^2}; \sqrt{R^2-y^2}\right]\right\}$$

$$A = A_1 \cup A_2 \cup A_3 \cup A_4$$

È facilmente dimostrabile che una funzione continua fino alla frontiera di un insieme regolare esclusa, ma comunque limitata fino alla frontiera, è integrabile su questo dominio.

Ma come si può calcolare questo integrale? Inanzitutto osserviamo che, dato che un dominio regolare è un'unione di più insiemi, per additività dell'intervallo di integrazione (proprietà che vale anche per integrali in più variabili), il problema si riconduce a trovare l'integrale di un insieme x-semplice o y-semplice. Prendiamo il secondo come esempio. Questo esempio lo definiamo quindi nel modo seguente:

$$\Omega = \left\{(x,y) \in \mathbb{R}^2 \ ; \ x \in [a,b] \ ; \ g_1(x) \leq y \leq g_2(x)\right\}$$

Inanzitutto, lasciamoci guidare dall'intuizione. Scomponiamo il volume del solido sotteso dalla superficie $z = f(x, y)$ sul dominio Ω in tante "lamine", con basi sul piano xy parallele all'asse x, e perpendicolari a questo piano. Possiamo dire che il volume del solido da ottenere sarà dato dalla somma delle superfici di ognuna di queste "lamine" moltiplicata per uno spessore "infinitesimo". Se chiamiamo la superficie di una lamina $A(x)$ (è funzione di x in quanto varia al variare della x), abbiamo che:

$$\iint_\Omega f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \sum A(x)\,\mathrm{d}x$$

Ma come sappiamo, la somma di infiniti termini infinitesimi è esprimibile come un integrale:

$$\iint_\Omega f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_a^b A(x)\,\mathrm{d}x$$

A questo punto dobbiamo trovare il valore di $A(x)$. Ma questa è semplicemente l'area sottesa dal grafico di una funzione. Questa funzione è quella che si ricava dalla $f(x, y)$ fissando una particolare x. L'integrale sarà dunque calcolato sulla y, ed i due estremi dipenderanno dalle funzione g_1 e g_2 che definiscono l'insieme:

$$A(x) = \int_{g_1(x)}^{g_2(x)} f(x,y)\,\mathrm{d}y$$

Ecco perchè A risulta una funzione di x: infatti la funzione viene solo integrata sulla y, e quindi solo la y sparisce dall'integrale, facendo rimanere il risultato un valore dipendente dalla x (che di solito appare sia nella funzione, sia negli estremi di integrazione).

Riassumendo tutto ciò in unico teorema:

Teorema 9.6.2 (Integrazione su domini y-semplici) *Dato un insieme y-semplice*

$$\Omega = \left\{(x,y) \in \mathbb{R}^2 \,;\, x \in [a,b] \,;\, g_1(x) \le y \le g_2(x)\right\}$$

ed una funzione $f : \mathbb{R}^2 \to \mathbb{R}$ continua all'interno di questo insieme e limitata, vale che:

$$\iint_\Omega f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_a^b \left[\int_{g_1(x)}^{g_2(x)} f(x,y)\,\mathrm{d}y\right]\,\mathrm{d}x$$

Il ragionamento per quanto riguarda gli insiemi x-semplici è ovviamente analogo.

Esempio 9.6.4 *Calcolare*

$$\iint_A (x^2y - e^{\sqrt{y}})\,\mathrm{d}x\,\mathrm{d}y$$

con

$$A = \left\{(x,y) \in \mathbb{R}^2 \;:\; 0 < x < 2 \,;\, 0 < y < x^2\right\}$$

A è facilmente riconoscibile come un insieme y-semplice, e quindi:

$$\iint_A (x^2y - e^{\sqrt{y}})\,\mathrm{d}x\,\mathrm{d}y = \int_0^2 \mathrm{d}x \int_0^{x^2} (x^2y - e^{\sqrt{y}})\,\mathrm{d}y = A$$

L'integrale dell'esponenziale si risolve facilmente ponendo $x = u^2$ e compiendo quindi prima un'integrazione per sostituzione e quindi un'altra per parti, ottenendo:

$$\int_0^2 \mathrm{d}x \left\{\left[x^2\frac{y^2}{2}\right]_0^{x^2} - 2(xe^x - e^x + 1)\right\} =$$

$$\int_0^2 \mathrm{d}x \left(\frac{1}{2}x^6 - 2(xe^x - e^x + 1)\right) =$$

Ricaviamo che:

$$\int_0^2 x^6\,\mathrm{d}x = \frac{1}{2}\left[\frac{x^7}{7}\right]_0^2 = \frac{2^6}{7} = \frac{64}{7}$$

$$\int_0^2 xe^x\,\mathrm{d}x = [xe^x]_0^2 - \int_0^2 e^x\,\mathrm{d}x = 2e^2 - e^2 + 1$$

$$\int_0^2 e^x\,\mathrm{d}x = e^2 - 1$$

$$\int_0^2 \mathrm{d}x = 2$$

E quindi, sommando i vari membri:

$$\iint_A (x^2y - e^{\sqrt{y}})\,\mathrm{d}x\,\mathrm{d}y = \frac{64}{7} - 2\left(2e^2 - e^2 + 1 - e^2 + 1 + 2\right) = \frac{64}{7} - 8$$

Come si può vedere, l'unica difficoltà degli integrali doppi risulta in una certa laboriosità di calcoli, che comunque sono già gli stessi incontrati nella risoluzione degli integrali semplici. Quello che abbiamo fatto è stato trasformare un integrale doppio in un *integrale iterato*, ovvero una serie di integrali l'uno contenuto nell'altro.

Già nell'esame degli integrali semplici, avevamo detto che ponevamo l'area di una superficie piana così descritta:

$$E = \left\{(x,y) \in \mathbb{R}^2 \ : \ a \leq x \leq b \ ; \ 0 \leq y \leq f(x)\right\}$$

Come:

$$A_E = \int_a^b f(x)\,\mathrm{d}x$$

Ora, attraverso gli integrali doppi, possiamo dare una definizione anche migliore di *area di una superficie piana regolare*:

$$A_\Omega = \iint_\Omega \mathrm{d}x\,\mathrm{d}y$$

Ovvero l'integrale della funzione costante $f(x,y) = 1$ sul dominio dato. Intuitivamente, questo consiste nel calcolare il volume di solido cilindrico C che ha per base il dominio da noi scelto e altezza 1. Ma questo volume è pari all'area di base per l'altezza, e quindi avremmo che:

$$\iint_\Omega \mathrm{d}x\,\mathrm{d}y = V_C = A_\Omega \cdot 1 = A_\Omega$$

Ecco quindi giustificata la regola data.

Precedentemente avevamo parlato dell'additività dell'integrale rispetto all'insieme di integrazione – ma l'integrale ha anche altre proprietà:

- Linearità dell'integrale:

$$\iint_\Omega [\lambda f(x,y) + \mu g(x,y)]\,\mathrm{d}x\,\mathrm{d}y = \lambda \iint_\Omega f(x,y)\,\mathrm{d}x\,\mathrm{d}y + \mu \iint_\Omega g(x,y)\,\mathrm{d}x\,\mathrm{d}y$$

- Monotonìa dell'integrale:

$$\forall \vec{x} \in \Omega : f(\vec{x}) \geq g(\vec{x}) \implies \iint_\Omega f(x,y)\,\mathrm{d}x\,\mathrm{d}y \geq \iint_\Omega g(x,y)\,\mathrm{d}x\,\mathrm{d}y$$

  e quindi in particolare:

$$\forall \vec{x} \in \Omega : f(\vec{x}) \geq 0 \implies \iint_\Omega f(x,y)\,\mathrm{d}x\,\mathrm{d}y \geq 0$$

- Additività rispetto al dominio di integrazione:

$$\iint_{\Omega_1 \cup \Omega_2} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \iint_{\Omega_1} f(x,y)\,\mathrm{d}x\,\mathrm{d}y + \iint_{\Omega_2} f(x,y)\,\mathrm{d}x\,\mathrm{d}y$$

- Condizioni di annullamento
  - Se il dominio ha area nulla, l'integrale è nullo.

$$A_\Omega = 0 \implies \iint_\Omega f(x,y)\,\mathrm{d}x\,\mathrm{d}y = 0$$

  - Se l'integrale di una funzione non-negativa è nullo su un dominio non nullo allora la funzione non solo è non negativa, ma nulla in tutto il dominio.

$$A_\Omega \neq 0 \ , \ \forall \vec{x} \in \Omega : f(\vec{x}) \geq 0 \ , \ \iint_\Omega f(x,y)\,\mathrm{d}x\,\mathrm{d}y = 0 \implies \forall \vec{x} \in \Omega : f(\vec{x}) = 0$$

  - Se su tutti i sottoinsiemi del del dominio di integrazione l'integrale è nullo, anche la funzione è nulla:

$$\forall D \subset \Omega : \iint_D f(x,y)\,\mathrm{d}x\,\mathrm{d}y = 0 \implies f(x,y) = 0$$

- **Teorema 9.6.3 (Teorema della media)** *Se Ω è connesso, allora:*

$$\exists \vec{x}_0 \in \Omega \ : \ \iint_\Omega f(x,y)\,\mathrm{d}x\,\mathrm{d}y = f(\vec{x}_0) A_\Omega$$

- Vale la disuguaglianza:

$$\left| \iint_\Omega f(x,y)\,\mathrm{d}x\,\mathrm{d}y \right| \leq \iint_\Omega |f(x,y)|\,\mathrm{d}x\,\mathrm{d}y \leq A_\Omega \cdot \max_\Omega |f|$$

Come annotazione, diciamo che un'applicazione pratica degli integrali doppi consiste nella ricerca dei baricentri e momenti d'inerzia. Le coordinate del baricentro di una figura piana Ω, con massa totale M e densità $\rho(x, y)$ è:

$$\left(\frac{1}{M}\iint_\Omega x\rho(x,y)\,\mathrm{d}x\,\mathrm{d}y\ ;\ \frac{1}{M}\iint_\Omega y\rho(x,y)\,\mathrm{d}x\,\mathrm{d}y\right)$$

E, utilizzando gli stessi simboli di prima, il suo momento d'inerzia intorno all'asse perpendicolare al piano xy:

$$I_\Omega = \iint_\Omega d^2(x,y)\rho(x,y)\,\mathrm{d}x\,\mathrm{d}y$$

Dove $d(x, y)$ restituisce la distanza di (x, y) dall'asse di rotazione.

9.6.3 Cambio di variabile e funzioni a valori vettoriali

Così come avevamo visto che, per quanto riguarda gli integrali semplici, il cambio di variabile (un risultato della composizione di funzioni) poteva portare ad integrali più facili ed immediati, anche nel campo degli integrali multipli possiamo semplificare la ricerca del risultato.

Cerchiamo inanzitutto, intuitivamente, di capire cosa significhi geometricamente un cambio di variabile. Per questo, prendiamo l'integrale

$$\iint_A \frac{1}{x^2+y^2+1}\,\mathrm{d}x\,\mathrm{d}y$$

dove l'insieme A, espresso in coordinate polari, risulta essere:

$$A = \{(\rho,\theta) \mid 0 \le \rho \le 1;\ 0 \le \theta \le 2\pi\}$$

Ed è quindi un rettangolo nel piano "polare", ed uno spicchio nel piano cartesiano. È ovvio che, a questo punto, la sostituzione in *tutto* l'integrale delle coordinate cartesiane in coordinate polari sarebbe, a quanto pare, una soluzione molto comoda. In questo modo avremmo un dominio rettangolare, ed una funzione da integrare alquanto semplice, infatti diventerebbe:

$$\frac{1}{x^2+y^2+1} = \frac{1}{1+\rho^2}$$

Abbiamo compiuto allora tutte le sostituzioni necessarie? A dire il vero no, perchè se si osserva bene, manca un ultimo fattore dentro l'integrale:

$$\mathrm{d}x\,\mathrm{d}y$$

Dal quale dobbiamo far apparire un $\mathrm{d}\rho\,\mathrm{d}\theta$. Pensiamo adesso al significato geometrico di quel termine. Il modo più immediato è quello di pensare ad un $\mathrm{d}x$ come ad un elemento infinitesimo sull'asse x, quindi una lunghezza infinitesima. Ed in effetti, dal significato di differenziale che conosciamo, sappiamo che esso approssima in modo perfetto una $f(x)$ solo per $x - x_0 \to 0$. Quindi, il prodotto di due differenziali risulterà essere una area infinitesima. Il prodotto della funzione per questa area infinitesima sarà quindi una sottilissima striscia verticale della funzione, e quindi un volume infinitesimo. L'integrale non fa quindi nient'altro che sommare queste aree infinitesime.

Allora, dopo aver convertito l'area di base e la funzione, rimane solo da convertire l'elemento di area infinitesimo $\mathrm{d}x\,\mathrm{d}y$ in $\mathrm{d}\rho\,\mathrm{d}\theta$. Prendiamo allora in un piano (cartesiano) un punto (ρ_0, θ_0), ed un punto a distanza infinitesima $(\rho_0 + \mathrm{d}\rho, \theta_0 + \mathrm{d}\theta)$. Se ora riusciamo...

TODO: *DA FINIRE - più che altro da capire...*

Ora, superata questa argomentazione alquanto intuitiva nel suo complesso, passiamo ad una formalizzazione dei concetti quì presentati. C'è da riconoscere inanzitutto che ci troviamo a trattare, per la prima volta, con delle *funzioni vettoriali a valori vettoriali*: infatti tentiamo di trasformare una funzione a più variabili con una loro sostituzione; ognuna di queste sostituzioni è una funzione, e quindi nel suo complesso abbiamo definito una funzione vettoriale a valori vettiali. Prediamo quindi una generica:

$$F : A \subseteq \mathbb{R}^n \to \mathbb{R}^p$$

Ora, per quello appena detto, una funzione vettoriale può essere vista come composizione di più funzioni reali, ovvero:

$$F(\vec{x}) = (F_1(\vec{x}), F_2(\vec{x}), \ldots, F_p(\vec{x}))$$

Dove $\vec{x} \in \mathbb{R}^q$. Ma adesso abbiamo che:

$$F_{1\le n\le p}(\vec{x}) \to \mathbb{R}$$

Con la stessa restrizione per $\vec{x}$. Quindi quello che ci troviamo a studiare è un insieme di p funzioni reali a valori vettoriali, che è il nostro ultimo campo di ricerca.

Questa volta per molte definizioni il nostro compito è facilitato dalle definizioni già date, ad esempio si dice che F è continua in $\vec{x}_0$ se e solo se le $F_1(\vec{x})$, $F_2(\vec{x})$, ..., $F_p(\vec{x})$ sono continue in $\vec{x}_0$, e lo stesso schema vale per la derivabilità, differenziabilità, ecc ...

Tuttavia qualcosa di perde, similmente a quanto era avvenuto in campo complesso (che in fondo non è altro che un particolare $\mathbb{R}^2$), ovvero la relazione d'ordine, e quindi il significato di massimo e minimo. Anche in questo caso è possibile dimostrare che qualsiasi ricerca di una relazione d'ordine in $\mathbb{R}^{n>1}$ è vana – dimostrazione che qui non inseriremo, tuttavia un semplice esempio mostrerà che il metodo più naturale per questa definizione non funziona.

Prendiamo infatti come sistema quello di dire che $\vec{a} \leq \vec{b}$ se e solo se $|a| \leq |b|$. Questa relazione, pur essendo transitiva e riflessiva, non è antisimmetrica. Prendiamo ad esempio i due punti $\vec{a} = (1, 0, 0)$ e $\vec{b} = (0, 1, 0)$. Per essi, abbiamo che $|a| = |b| = 1$, quindi $\vec{a} \leq \vec{b}$, ma anche $\vec{b} \leq \vec{a}$. Tuttavia, *non* è vero che $\vec{a} = \vec{b}$!

Quindi, messo da parte questo problema, un altro che si può essere utile risolvere è quello di verificare se una trasformazione è invertibile o meno. Ora, se F è differenziabile, prendiamo la seguente matrice, chiamata *matrice Jacobiana di F in $\vec{x}_0$*:

$$J_F(\vec{x}_0) = \begin{bmatrix} \frac{\partial F_1}{\partial x_1}(\vec{x}_0) & \frac{\partial F_1}{\partial x_2}(\vec{x}_0) & \dots & \frac{\partial F_1}{\partial x_q}(\vec{x}_0) \\ \frac{\partial F_2}{\partial x_1}(\vec{x}_0) & \frac{\partial F_2}{\partial x_2}(\vec{x}_0) & \dots & \frac{\partial F_2}{\partial x_q}(\vec{x}_0) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial F_p}{\partial x_1}(\vec{x}_0) & \frac{\partial F_p}{\partial x_2}(\vec{x}_0) & \dots & \frac{\partial F_p}{\partial x_q}(\vec{x}_0) \end{bmatrix}$$

In pratica, l'elemento generico alla riga i e colonna j è:

$$a_{ij} = \frac{\partial F_i}{\partial x_j}(\vec{x}_0)$$

Con $1 \leq i \leq p$ e $1 \leq j \leq q$, ovvero la matrice Jacobiana è una matrice $p \times q$.

Esempio 9.6.5 *Trovare il valore della matrice Jacobiana della funzione*

$$F : \mathbb{R}^3 \to \mathbb{R}^2; \quad F(x, y, z) = (e^x \cos y \sin z, e^x \sin y \sin z)$$

nel punto $(0, 2\pi, \pi/2)$.

Abbiamo quindi le due funzioni:

$$F_1(x, y, z) = e^x \cos y \sin z$$

e

$$F_2(x, y, z) = e^x \sin y \sin z$$

Calcoliamo quindi le varie derivate parziali (dove ovviamente avremo $x_1 = x$, $x_2 = ye$ e $x_3 = z$):

$$\frac{\partial F_1}{\partial x} = e^x \cos y \sin z$$

$$\frac{\partial F_1}{\partial y} = -e^x \sin y \sin z$$

$$\frac{\partial F_1}{\partial z} = e^x \cos y \cos z$$

Passiamo ora ad F_2:

$$\frac{\partial F_2}{\partial x} = e^x \sin y \sin z$$

$$\frac{\partial F_2}{\partial y} = e^x \cos y \sin z$$

$$\frac{\partial F_2}{\partial z} = e^x \sin y \cos z$$

A questo punto siamo pronti per costruire la matrice Jacobiana generica:

$$J_F(x, y, z) = \begin{bmatrix} e^x \cos y \sin z & -e^x \sin y \sin z & e^x \cos y \cos z \\ e^x \sin y \sin z & e^x \cos y \sin z & e^x \sin y \cos z \end{bmatrix}$$

E passiamo ora a trovarla nel punto $(0, 2\pi, \pi/2)$*:*

$$J_F(0, 2\pi, \pi/2) = \begin{bmatrix} 1 \cdot 1 \cdot 1 & -1 \cdot 0 \cdot \dots & 1 \cdot 1 \cdot 0 \\ 1 \cdot 0 \cdot \dots & 1 \cdot 1 \cdot 1 & 1 \cdot 0 \cdot \dots \end{bmatrix}$$

Ovvero:

$$J_F(0, 2\pi, \pi/2) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

Ora vedremo l'importanza della matrice Jacobiana.

Se abbiamo una funzione $f : \mathbb{R}^q \to \mathbb{R}$, dire che f è differenziabile sappiamo che implica che:

$$f(\vec{x}_0 + \vec{h}) = f(\vec{x}_0) + \overrightarrow{\nabla} f(\vec{x}_0)\vec{h} + o(|\vec{h}|)$$

Quindi, se ora torniamo alla nostra $F : \mathbb{R}^q \to \mathbb{R}^p$, abbiamo che, per ognuna delle sue componenti:

$$F_j(\vec{x}_0 + \vec{h}) = F_j(\vec{x}_0) + \overrightarrow{\nabla} F_j(\vec{x}_0)\vec{h} + o(|\vec{h}|)$$

Ora, se volessimo considerare tutte le varie componenti assieme, ovvero l'intera F, avremo che il primo addendo diventa semplicemente $F(\vec{x}_0)$ (poichè ogni componente di quest'ultimo vettore corrisponde con i vari $F_j(\vec{x}_0)$), ed i vari $o(|\vec{h}|)$ diventeranno un unico $o(|\vec{h}|)$, questa volta inteso come:

$$\lim_{\vec{h}\to\vec{0}} \frac{o(|\vec{h}|)}{|\vec{h}|} = \vec{0}$$

(ovvero: la funzione considerata questa volta è vettoriale a valori vettoriali), anche questa volta per semplice corrispondenza tra componente e componente. A questo punto consideriamo il secondo addendo. Ognuno di essi è, in definitiva, il prodotto tra la riga j-esima della matrice Jacobiana e il vettore colonna $\vec{h}$, e quindi può essere riassunta proprio con $J_F(\vec{x}_0) \cdot \vec{h}$. La formula che cercavamo risulta quindi essere:

$$F(\vec{x}_0 + \vec{h}) = F(\vec{x}_0) + J_F(\vec{x}_0) \cdot \vec{h} + o(|\vec{h}|)$$

O anche, come al solito:

$$F(\vec{x}_0 + \vec{h}) - F(\vec{x}_0) = J_F(\vec{x}_0) \cdot \vec{h} + o(|\vec{h}|)$$

Il che significa che, presa la differenza tra i valori che la F assume se calcolata in punti distanti $\vec{h}$, abbiamo che essa è uguale ad una *trasformazione lineare* del vettore $\vec{h}$, a meno di un errore che è un o piccolo di $|\vec{h}|$. Vale anche qui, quindi, lo stesso discorso fatto a suo tempo per la formula di Taylor al prim'ordine: non solo decresce il valore dell'errore all'avvicinarsi di $|\vec{h}|$ a zero, ma decresce anche il rapporto tra l'errore ed $|\vec{h}|$, che in condizioni normali creerebbe una forma indeterminata del tipo $0/0$.

Esempio 9.6.6 *Approssimare la funzione*

$$F : \mathbb{R}^2 \to \mathbb{R}^2; \quad F(x, y) = (x + y^2, y - x^2)$$

nel punto $\vec{0}$ e verificare che, effettivamente, l'errore commesso è un o piccolo di $|\vec{h}|$.

F_1 ed F_2 sono derivabili infinite volte in modo continuo (appartengono quindi a $C^\infty(A)$), e sono quindi differenziabili e per esse possiamo compiere l'approssimazione al prim'ordine.

Calcoliamo quindi la matrice Jacobiana:

$$J_F(x, y) = \begin{bmatrix} 1 & 2y \\ -2x & 1 \end{bmatrix}$$

E quindi:

$$J_F(0, 0) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2$$

E, posto $\vec{h} = (h_1, h_2)$, abbiamo che:

$$F(\vec{x}_0 + \vec{h}) = F(\vec{0} + \vec{h}) = F(h_1, h_2) = (h_1 - h_2^2, h_2 - h_1^2)$$

e

$$F(\vec{x}_0) = F(0, 0) = (0, 0)$$

Quindi, dato che abbiamo visto che la matrice Jacobiana nel nostro caso equivale alla matrice identità:

$$J_F(0, 0) \cdot \vec{h} = \vec{h}$$

Riassumento il tutto, troviamo quindi che:

$$F(\vec{x}_0 + \vec{h}) = (0, 0) + (h_1, h_2) + \vec{\varepsilon}(h_1, h_2)$$

Dove $\vec{\varepsilon}$ è il nostro errore. Ora, per i risultati trovati, è facile ricavare che:

$$\vec{\varepsilon}(h_1, h_2) = (h_2^2, -h_1^2)$$

Ed in effetti ora possiamo verificare che questo è un o piccolo di $|\vec{h}|$, infatti:

$$\frac{\vec{\varepsilon}}{|\vec{h}|} = \frac{(h_2^2, -h_1^2)}{\sqrt{h_1^2 + h_2^2}} = \left(\frac{h_2^2}{\sqrt{h_1^2 + h_2^2}}, -\frac{h_1^2}{\sqrt{h_1^2 + h_2^2}} \right)$$

Che è facile dimostrare che tende a 0, per $\vec{h} \to \vec{0}$. Quindi, effettivamente, abbiamo che, per questa F:

$$F(\vec{x}_0 + \vec{h}) = F(\vec{x}_0) + J_F(\vec{x}_0) \cdot \vec{h} + o(|\vec{h}|)$$

In $\mathbb{R} \to \mathbb{R}$ avevamo che se una funzione f aveva $f' > 0$, ovvero era strettamente crescente, allora era sicuramente invertibile. Si potrebbe quindi fare la congettura che se

$$F : \mathbb{R}^q \to \mathbb{R}^q,\ q > 1$$

e

$$\forall \vec{x} : \det J_F(\vec{x}) > 0$$

ovvero la matrice Jacobiana è invertibile (poichè il suo determinante non è mai nullo), allora la funzione F è a sua volta invertibile. In verità questo *non* accade, ed è facile mostrarlo con qualche esempio.

Esempio 9.6.7 *Si prenda la funzione*

$$F : \mathbb{R}^2 \to \mathbb{R}^2; \quad F(x, y) = (e^x \cos y, e^x \sin y)$$

E si dimostri che non vale la congettura proposta.

F è ovviamente differenziabile, quindi anche continua e derivabile, e possiamo calcolare il suo Jacobiano:

$$J_F(x, y) = \begin{bmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{bmatrix}$$

Ed il suo determinante vale:

$$\det J_f(x, y) = e^{2x} \cos^2 y + e^{2x} \sin^y = e^2 x \neq 0, \forall (x, y)$$

Tuttavia, la funzione è chiaramente non invertibile, infatti abbiamo che:

$$\forall k \in \mathbb{Z} \ :\ F(x, y) = F(x, y + 2k\pi)$$

E quindi, dato che F non è iniettiva, non è neanche biunivoca e quindi invertibile.

Tuttavia, sotto conclusioni meno ampie di quelle da noi proposte, si scopre che la congettura rimane comunque valida, infatti vale il seguente: **TODO**: *Il teorema parla dell'iniettività o della biunivocità?*

Teorema 9.6.4 (Invertibilità delle funzione vettoriali a valori vettoriali) *Data $F : \mathbb{R}^q \to \mathbb{R}^q$, se $\det J_F(\vec{x}_0) \neq 0$, allora esiste un intorno I_{x_0} (abbastanza piccolo) di $\vec{x}_0$ nel quale $F|I_{x_0}$ è invertibile, con inversa differenziabile, e inoltre:*

$$J_{F^{-1}}(F(\vec{x})) = (J_F(\vec{x}))^{-1}$$

L'ultima conseguenza è poi simile alla derivata di una funzione inversa in $\mathbb{R} \to \mathbb{R}$:

$$(f^{-1})'(x) = \frac{1}{f'(f^{-1}(x))}$$

Esempio 9.6.8 *Verificare il teorema dato per l'esempio precedente*

Noi sapevamo che il problema principale della funzione precedente era la sua periodicità di 2π, che impediva ad essa di essere invertibile. Quindi, basterà limitarci a prendere un intervallo più breve, e quindi un dominio più ristretto (che sarebbe poi l'intorno del teorema):

$$\Omega = \left\{(x, y) \in \mathbb{R}^2 \mid -\pi < y < \pi\right\}$$

Questo insieme Ω sarà quindi la fascia orizzontale di altezza 2π ed estensione sulla x infinita.

Effettivamente, la funzione così trovata è iniettiva. La condizione che dobbiamo quindi dimostrare è:

$$(x_1, y_1) \neq (x_2, y_2) \Longrightarrow F(x_1, y_1) \neq F(x_2, y_2)$$

Ma questo equivale a dire che:

$$F(x_1, y_1) = F(x_2, y_2) \Longrightarrow (x_1, y_1) = (x_2, y_2)$$

Ora, se $F(x_1, y_1) = F(x_2, y_2)$, allora sarà anche che $|F(x_1, y_1)| = |F(x_2, y_2)|$, e quindi:

$$\sqrt{e^{2x_1} \cos^2 y_1 + e^{2x_1} \sin^2 y_1} = \sqrt{e^{2x_2} \cos^2 y_2 + e^{2x_2} \sin^2 y_2}$$

Al che ricaviamo

$$e^{x_1} = e^{x_2}$$

Ovvero:

$$x_1 = x_2$$

E così abbiamo trovato la condizione sulle x. Ora quindi l'uguaglianza delle due funzioni può essere riscritta come:

$$(e_1^x \cos y_1, e_1^x \sin y_1) = (e_1^x \cos y_2, e_1^x \sin y_2)$$

(notare la sostituzione $x_1 = x_2$), ovvero:

$$\begin{cases} \cos y_1 = \cos y_2 \\ \sin y_1 = \sin y_2 \end{cases}$$

Ora, nella circonferenza trigonometrica abbiamo sempre e solo un punto nel quale seno e coseno assumono un particolare valore. Quindi, otteniamo che:

$$\forall k \in \mathbb{Z} \ : \ y_1 = y_2 + 2k\pi$$

Ma dato che abbiamo posto la limitazione $-\pi < y < \pi$, otteniamo:

$$y_1 = y_2$$

Quindi ecco dimostrata l'iniettività della funzione data. A questo punto dobbiamo dimostrarne la suriettività, e questo sarà possibile trovando il codominio della funzione:

$$F(\Omega)$$

Questo si rivela sempre alquanto difficile per funzioni a valori vettoriali.

Intanto, sicuramente avremo che:

$$\nexists \vec{x} \ ; \ F(\vec{x}) = (0,0)$$

Poichè la funzione esponenziale non è mai nulla, e le funzioni seno e coseno non si annullano mai contemporaneamente. Ora, facciamo la congettura che:

$$F(\Omega) = \mathbb{R}^2 - \{(0,0)\}$$

Per dimostrarlo occorrerà prendere un punto $(u,v) \neq (0,0)$, e dimostrare che $\exists (x,y) \ : \ F(x,y) = (u,v)$. Quindi occorre risolvere il sistema:

$$\begin{cases} e^x \cos y = u \\ e^x \sin y = v \end{cases}$$

A questo punto sommiamo i quadrati delle due espressioni membro a membro:

$$e^{2x} \cos^2 y + e^{2x} \sin^2 y = u^2 + v^2$$

Ovvero

$$e^{2x} = u^2 + v^2$$

Da cui si ricava:

$$x = \frac{1}{2} \ln(u^2 + v^2)$$

E, riprendendo le equazioni iniziali, troviamo:

$$\begin{cases} \cos y = \dfrac{u}{\sqrt{u^2 + v^2}} \\ \\ \sin y = \dfrac{v}{\sqrt{u^2 + v^2}} \end{cases}$$

Sicuramente sarà possibile sempre trovare un punto y nel dominio proposto che soddisfi queste equazioni (basta osservare che $\cos^2 y + \sin^2 y = 1$, e quindi sono punti della circonferenza trigonometrica). Ecco allora che effettivamente abbiamo:

$$F(\Omega) = \mathbb{R}^2 \left\{ (x,y) \in \mathbb{R}^2 \mid x \leq 0, y = 0 \right\}$$

***TODO**: Perche' $x \leq 0$??*

Adesso siamo di nuovo pronti a tornare al nostro argomento principale, ovvero la ricerca di una approssimazione delle aree sotto cambio di variabile. Prendiamo quindi una

$$F : \mathbb{R}^2 \to \mathbb{R}^2$$

differenziabile. Per semplicità prendiamo che

$$F(0,0) = (0,0)$$

Prendiamo ora un quadrato con un vertice in $(0,0)$ e lati Δx e Δy lungo i due assi. Per la precisione prendiamo i due vettori:

$$\vec{\Delta} x = (\Delta x, 0)$$

$$\vec{\Delta} y = (0, \Delta y)$$

Prendiamo ora questo dominio quadrato e lo trasformiamo secondo la funzione data. Tuttavia, possiamo considerare prima un'approssimazione che ci sarà molto utile, infatti sappiamo, per quello appena detto, che, a meno di un infinitesimo nell'ordine di $o(|\vec{h}|)$, abbiamo:

$$F(\vec{x}_0 + \vec{h}) = F(\vec{x}_0) + J_F(\vec{x}_0) \cdot \vec{h}$$

Se qui prendiamo $\vec{x}_0 = (0,0)$ ed $\vec{h} = \vec{\Delta}x$ (oppure $\vec{h} = \vec{\Delta}y$ per l'altro lato), abbiamo:

$$F(\vec{\Delta}x) = F(0,0) + J_F(\vec{0}) \cdot \vec{\Delta}x$$

e

$$F(\vec{\Delta}y) = F(0,0) + J_F(\vec{0}) \cdot \vec{\Delta}y$$

Dove abbiamo:

$$J_F(\vec{0}) = \begin{bmatrix} \frac{\partial F_1}{\partial x}(0,0) & \frac{\partial F_1}{\partial y}(0,0) \\ \frac{\partial F_2}{\partial x}(0,0) & \frac{\partial F_2}{\partial y}(0,0) \end{bmatrix}$$

Approssimiamo quindi la trasformazione del $\vec{\Delta}x$ con

$$\vec{a} = F(\vec{\Delta}x) \approx J_F(\vec{0}) \cdot (\Delta x, 0) = \begin{bmatrix} \frac{\partial F_1}{\partial x}(0,0)\Delta x \\ \frac{\partial F_2}{\partial x}(0,0)\Delta x \end{bmatrix}$$

E, analogamente:

$$\vec{b} = F(\vec{\Delta}y) \approx J_F(\vec{0}) \cdot (0, \Delta y) = \begin{bmatrix} \frac{\partial F_1}{\partial y}(0,0)\Delta y \\ \frac{\partial F_2}{\partial y}(0,0)\Delta y \end{bmatrix}$$

Ora, l'area del quadrato di partenza aveva per area, evidentemente:

$$A_Q = \Delta x \cdot \Delta y$$

Adresso invece abbiamo un parallelogramma di cui calcolare l'area. Come sappiamo, la sua area equivale a:

$$A_{Q'} = |\vec{a} \cdot \vec{b}| = \left| \begin{bmatrix} \vec{i} & \vec{j} & \vec{k} \\ \frac{\partial F_1}{\partial x}\Delta x & \frac{\partial F_2}{\partial x}\Delta x & 0 \\ \frac{\partial F_1}{\partial y}\Delta y & \frac{\partial F_2}{\partial y}\Delta y & 0 \end{bmatrix} \right| =$$

Semplificando, otteniamo:

$$\frac{\partial F_1}{\partial x}(0,0)\frac{\partial F_2}{\partial y}(0,0)\Delta x\Delta y - \frac{\partial F_2}{\partial x}(0,0)\frac{\partial F_1}{\partial y}(0,0)\Delta x\Delta y$$

Che in realtà non è nient'altro che:

$$\Delta x\Delta y |\det J_F(0,0)| = A_Q |\det J_F(0,0)|$$

E da questa formula, si può generalizzare in:

$$A_{F(Q)} \approx A_Q |\det J_F(\overline{x}, \overline{y})|$$

E similmente per i volumi:

$$V_{F(P)} \approx V_P |\det J_F(\overline{x}, \overline{y}, \overline{z})|$$

9.6.4 Coordinate polari, cilindriche, sferiche

Alcune funzioni di trasformazione sono particolarmente interessanti.

La prima è quella delle coordinate polari, definite da:

Appendice A

GNU Free Documentation License

Version 1.1, March 2000

Preamble

The purpose of this License is to make a manual, textbook, or other written document "free" in the sense of freedom: to assure everyone the effective freedom to copy and redistribute it, with or without modifying it, either commercially or noncommercially. Secondarily, this License preserves for the author and publisher a way to get credit for their work, while not being considered responsible for modifications made by others.

This License is a kind of "copyleft", which means that derivative works of the document must themselves be free in the same sense. It complements the GNU General Public License, which is a copyleft license designed for free software.

We have designed this License in order to use it for manuals for free software, because free software needs free documentation: a free program should come with manuals providing the same freedoms that the software does. But this License is not limited to software manuals; it can be used for any textual work, regardless of subject matter or whether it is published as a printed book. We recommend this License principally for works whose purpose is instruction or reference.

A.1 Applicability and Definitions

This License applies to any manual or other work that contains a notice placed by the copyright holder saying it can be distributed under the terms of this License. The "Document", below, refers to any such manual or work. Any member of the public is a licensee, and is addressed as "you".

A "Modified Version" of the Document means any work containing the Document or a portion of it, either copied verbatim, or with modifications and/or translated into another language.

A "Secondary Section" is a named appendix or a front-matter section of the Document that deals exclusively with the relationship of the publishers or authors of the Document to the Document's overall subject (or to related matters) and contains nothing that could fall directly within that overall subject. (For example, if the Document is in part a textbook of mathematics, a Secondary Section may not explain any mathematics.) The relationship could be a matter of historical connection with the subject or with related matters, or of legal, commercial, philosophical, ethical or political position regarding them.

The "Invariant Sections" are certain Secondary Sections whose titles are designated, as being those of Invariant Sections, in the notice that says that the Document is released under this License.

The "Cover Texts" are certain short passages of text that are listed, as Front-Cover Texts or Back-Cover Texts, in the notice that says that the Document is released under this License.

A "Transparent" copy of the Document means a machine-readable copy, represented in a format whose specification is available to the general public, whose contents can be viewed and edited directly and straightforwardly with generic text editors or (for images composed of pixels) generic paint programs or (for drawings) some widely available drawing editor, and that is suitable for input to text formatters or for automatic translation to a variety of formats suitable for input to text formatters. A copy made in an otherwise Transparent file format whose markup has been designed to thwart or discourage subsequent modification by readers is not Transparent. A copy that is not "Transparent" is called "Opaque".

Examples of suitable formats for Transparent copies include plain ASCII without markup, Texinfo input format, LaTeX input format, SGML or XML using a publicly available DTD, and standard-conforming simple HTML designed for human modification. Opaque formats include PostScript, PDF, proprietary formats that can be read and edited only by proprietary word processors, SGML or XML for which the DTD and/or processing tools are not generally available, and the machine-generated HTML produced by some word processors for output purposes only.

The "Title Page" means, for a printed book, the title page itself, plus such following pages as are needed to hold, legibly, the material this License requires to appear in the title page. For works in formats which do not have any title page as such, "Title Page" means the text near the most prominent appearance of the work's title, preceding the beginning of the body of the text.

A.2 Verbatim Copying

You may copy and distribute the Document in any medium, either commercially or noncommercially, provided that this License, the copyright notices, and the license notice saying this License applies to the Document are reproduced in all copies, and that you add no other conditions whatsoever to those of this License. You may not use technical measures to obstruct or control the reading or further copying of the copies you make or distribute. However, you may accept compensation in exchange for copies. If you distribute a large enough number of copies you must also follow the conditions in section 3.

You may also lend copies, under the same conditions stated above, and you may publicly display copies.

A.3 Copying in Quantity

If you publish printed copies of the Document numbering more than 100, and the Document's license notice requires Cover Texts, you must enclose the copies in covers that carry, clearly and legibly, all these Cover Texts: Front-Cover Texts on the front cover, and Back-Cover Texts on the back cover. Both covers must also clearly and legibly identify you as the publisher of these copies. The front cover must present the full title with all words of the title equally prominent and visible. You may add other material on the covers in addition. Copying with changes limited to the covers, as long as they preserve the title of the Document and satisfy these conditions, can be treated as verbatim copying in other respects.

If the required texts for either cover are too voluminous to fit legibly, you should put the first ones listed (as many as fit reasonably) on the actual cover, and continue the rest onto adjacent pages.

If you publish or distribute Opaque copies of the Document numbering more than 100, you must either include a machine-readable Transparent copy along with each Opaque copy, or state in or with each Opaque copy a publicly-accessible computer-network location containing a complete Transparent copy of the Document, free of added material, which the general network-using public has access to download anonymously at no charge using public-standard network protocols. If you use the latter option, you must take reasonably prudent steps, when you begin distribution of Opaque copies in quantity, to ensure that this Transparent copy will remain thus accessible at the stated location until at least one year after the last time you distribute an Opaque copy (directly or through your agents or retailers) of that edition to the public.

It is requested, but not required, that you contact the authors of the Document well before redistributing any large number of copies, to give them a chance to provide you with an updated version of the Document.

A.4 Modifications

You may copy and distribute a Modified Version of the Document under the conditions of sections 2 and 3 above, provided that you release the Modified Version under precisely this License, with the Modified Version filling the role of the Document, thus licensing distribution and modification of the Modified Version to whoever possesses a copy of it. In addition, you must do these things in the Modified Version:

- Use in the Title Page (and on the covers, if any) a title distinct from that of the Document, and from those of previous versions (which should, if there were any, be listed in the History section of the Document). You may use the same title as a previous version if the original publisher of that version gives permission.
- List on the Title Page, as authors, one or more persons or entities responsible for authorship of the modifications in the Modified Version, together with at least five of the principal authors of the Document (all of its principal authors, if it has less than five).
- State on the Title page the name of the publisher of the Modified Version, as the publisher.
- Preserve all the copyright notices of the Document.
- Add an appropriate copyright notice for your modifications adjacent to the other copyright notices.
- Include, immediately after the copyright notices, a license notice giving the public permission to use the Modified Version under the terms of this License, in the form shown in the Addendum below.
- Preserve in that license notice the full lists of Invariant Sections and required Cover Texts given in the Document's license notice.
- Include an unaltered copy of this License.
- Preserve the section entitled "History", and its title, and add to it an item stating at least the title, year, new authors, and publisher of the Modified Version as given on the Title Page. If there is no section entitled "History" in the Document, create one stating the title, year, authors, and publisher of the Document as given on its Title Page, then add an item describing the Modified Version as stated in the previous sentence.

- Preserve the network location, if any, given in the Document for public access to a Transparent copy of the Document, and likewise the network locations given in the Document for previous versions it was based on. These may be placed in the "History" section. You may omit a network location for a work that was published at least four years before the Document itself, or if the original publisher of the version it refers to gives permission.
- In any section entitled "Acknowledgements" or "Dedications", preserve the section's title, and preserve in the section all the substance and tone of each of the contributor acknowledgements and/or dedications given therein.
- Preserve all the Invariant Sections of the Document, unaltered in their text and in their titles. Section numbers or the equivalent are not considered part of the section titles.
- Delete any section entitled "Endorsements". Such a section may not be included in the Modified Version.
- Do not retitle any existing section as "Endorsements" or to conflict in title with any Invariant Section.

If the Modified Version includes new front-matter sections or appendices that qualify as Secondary Sections and contain no material copied from the Document, you may at your option designate some or all of these sections as invariant. To do this, add their titles to the list of Invariant Sections in the Modified Version's license notice. These titles must be distinct from any other section titles.

You may add a section entitled "Endorsements", provided it contains nothing but endorsements of your Modified Version by various parties – for example, statements of peer review or that the text has been approved by an organization as the authoritative definition of a standard.

You may add a passage of up to five words as a Front-Cover Text, and a passage of up to 25 words as a Back-Cover Text, to the end of the list of Cover Texts in the Modified Version. Only one passage of Front-Cover Text and one of Back-Cover Text may be added by (or through arrangements made by) any one entity. If the Document already includes a cover text for the same cover, previously added by you or by arrangement made by the same entity you are acting on behalf of, you may not add another; but you may replace the old one, on explicit permission from the previous publisher that added the old one.

The author(s) and publisher(s) of the Document do not by this License give permission to use their names for publicity for or to assert or imply endorsement of any Modified Version.

A.5 Combining Documents

You may combine the Document with other documents released under this License, under the terms defined in section 4 above for modified versions, provided that you include in the combination all of the Invariant Sections of all of the original documents, unmodified, and list them all as Invariant Sections of your combined work in its license notice.

The combined work need only contain one copy of this License, and multiple identical Invariant Sections may be replaced with a single copy. If there are multiple Invariant Sections with the same name but different contents, make the title of each such section unique by adding at the end of it, in parentheses, the name of the original author or publisher of that section if known, or else a unique number. Make the same adjustment to the section titles in the list of Invariant Sections in the license notice of the combined work.

In the combination, you must combine any sections entitled "History" in the various original documents, forming one section entitled "History"; likewise combine any sections entitled "Acknowledgements", and any sections entitled "Dedications". You must delete all sections entitled "Endorsements."

A.6 Collections of Documents

You may make a collection consisting of the Document and other documents released under this License, and replace the individual copies of this License in the various documents with a single copy that is included in the collection, provided that you follow the rules of this License for verbatim copying of each of the documents in all other respects.

You may extract a single document from such a collection, and distribute it individually under this License, provided you insert a copy of this License into the extracted document, and follow this License in all other respects regarding verbatim copying of that document.

A.7 Aggregation With Independent Works

A compilation of the Document or its derivatives with other separate and independent documents or works, in or on a volume of a storage or distribution medium, does not as a whole count as a Modified Version of the Document, provided no compilation copyright is claimed for the compilation. Such a compilation is called an "aggregate", and this License does not apply to the other self-contained works thus compiled with the Document, on account of their being thus compiled, if they are not themselves derivative works of the Document.

If the Cover Text requirement of section 3 is applicable to these copies of the Document, then if the Document is less than one quarter of the entire aggregate, the Document's Cover Texts may be placed on covers that surround only the Document within the aggregate. Otherwise they must appear on covers around the whole aggregate.

A.8 Translation

Translation is considered a kind of modification, so you may distribute translations of the Document under the terms of section 4. Replacing Invariant Sections with translations requires special permission from their copyright holders, but you may include translations of some or all Invariant Sections in addition to the original versions of these Invariant Sections. You may include a translation of this License provided that you also include the original English version of this License. In case of a disagreement between the translation and the original English version of this License, the original English version will prevail.

A.9 Termination

You may not copy, modify, sublicense, or distribute the Document except as expressly provided for under this License. Any other attempt to copy, modify, sublicense or distribute the Document is void, and will automatically terminate your rights under this License. However, parties who have received copies, or rights, from you under this License will not have their licenses terminated so long as such parties remain in full compliance.

A.10 Future Revisions of This License

The Free Software Foundation may publish new, revised versions of the GNU Free Documentation License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns. See http://www.gnu.org/copyleft/.

Each version of the License is given a distinguishing version number. If the Document specifies that a particular numbered version of this License or any later version applies to it, you have the option of following the terms and conditions either of that specified version or of any later version that has been published (not as a draft) by the Free Software Foundation. If the Document does not specify a version number of this License, you may choose any version ever published (not as a draft) by the Free Software Foundation.

ADDENDUM: How to use this License for your documents

To use this License in a document you have written, include a copy of the License in the document and put the following copyright and license notices just after the title page:

> Copyright © YEAR YOUR NAME. Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.1 or any later version published by the Free Software Foundation; with the Invariant Sections being LIST THEIR TITLES, with the Front-Cover Texts being LIST, and with the Back-Cover Texts being LIST. A copy of the license is included in the section entitled “GNU Free Documentation License”.

If you have no Invariant Sections, write “with no Invariant Sections” instead of saying which ones are invariant. If you have no Front-Cover Texts, write “no Front-Cover Texts” instead of “Front-Cover Texts being LIST”; likewise for Back-Cover Texts.

If your document contains nontrivial examples of program code, we recommend releasing these examples in parallel under your choice of free software license, such as the GNU General Public License, to permit their use in free software.

Indice analitico