

Appunti di Analisi Matematica

Dr. Luca Granieri

2008

Indice

Introduzione	i
1 Continuità per le funzioni di una variabile reale	1
1.1 Continuità	2
1.2 Successioni	3
2 Integrazione	7
2.1 Definizione di integrale secondo Riemann	7
2.2 Cenni alla misura e all'integrale di Lebesgue	13
2.3 Il Teorema fondamentale del calcolo integrale	15
2.4 Alcuni metodi di integrazione	17
2.4.1 Integrazione per parti	17
2.4.2 Integrazione per sostituzione	18
2.4.3 Integrazione di funzioni razionali	20
2.5 Integrali generalizzati	22
2.6 Esercizi relativi al capitolo 1	26
3 Calcolo differenziale per funzioni di più variabili	27
3.1 Cenni di topologia di \mathbb{R}^N	27
3.2 Richiami di algebra lineare	30
3.2.1 Matrici e applicazioni lineari	31
3.3 Sistemi lineari	35
3.3.1 Il metodo di eliminazione di Gauss	36
3.4 Derivate parziali e differenziabilità	39
3.4.1 Derivate parziali	39
3.4.2 Derivate successive e Teorema di Schwarz	42
3.4.3 Differenziale di funzioni di più variabili	43

Capitolo 1

Continuità per le funzioni di una variabile reale

Ricordiamo le proprietà principali degli estremi superiori e inferiori che lasciamo come utile esercizio.

Esercizio 1. Sia $A \subset \mathbb{R}$ e sia y un maggiorante per A . Le seguenti affermazioni sono equivalenti:

1. $y = \sup A$,
2. $\forall z < y \exists x \in A$ t.c. $z < x$,
3. $\forall \varepsilon > 0 \exists x \in A$ t.c. $y < x + \varepsilon$.

Per l'estremo inferiore vale un risultato analogo:

Esercizio 2. Sia $A, B \subset \mathbb{R}$ e sia y un minorante per A . Le seguenti affermazioni sono equivalenti:

1. $y = \inf A$,
2. $\forall y < z \exists x \in A$ t.c. $x < z$,
3. $\forall \varepsilon > 0 \exists x \in A$ t.c. $x < y + \varepsilon$.

Sarà utile anche il seguente

Esercizio 3. Se $A, B \in \mathbb{R}$ allora si verificano:

1. $\sup(A + B) \leq \sup A + \sup B$,
2. $\inf A + \inf B \leq \inf(A + B)$,
3. $A \subset B \Rightarrow \inf B \leq \inf A$,
4. $A \subset B \Rightarrow \sup A \leq \sup B$,
5. se $\lambda > 0$, allora $\sup(\lambda A) = \lambda \sup A$, $\inf(\lambda A) = \lambda \inf A$,
6. se $\lambda < 0$, allora $\sup(\lambda A) = \lambda \inf A$, $\inf(\lambda A) = \lambda \sup A$.

1.1 Continuità

Definizione 4 (Continuità). Una funzione $f : A \rightarrow \mathbb{R}$ si dice continua in un punto $x_0 \in A$ se

$$\forall \varepsilon > 0 \exists \delta > 0 \text{ tale che } \forall x \in A : |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon.$$

f si dice continua in A se lo è in tutti i punti di A .

Esempio 5. $f(x) = x^2$ è continua in \mathbb{R} . Fissati $x_0 \in \mathbb{R}$ ed $\varepsilon > 0$ valutiamo che

$$|f(x) - f(x_0)| = |x^2 - x_0^2| = |x - x_0| \cdot |x + x_0| \leq |x - x_0|(|x - x_0| + 2|x_0|).$$

Allora si ha

$$|x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \delta^2 + 2\delta|x_0|.$$

Quindi basta prendere δ soluzione positiva dell'equazione $\delta^2 + 2\delta|x_0| - \varepsilon$ per verificare la definizione di continuità.

Osserviamo che nell'esempio precedente la scelta di δ dipende anche dal punto x_0 considerato oltre che da ε . Le funzioni per cui tale scelta può essere fatta indipendentemente dalla scelta del punto x_0 si dice uniformemente continue.

Definizione 6 (Continuità uniforme). Una funzione $f : A \rightarrow \mathbb{R}$ si dice uniformemente continua in A se

$$\forall \varepsilon > 0 \exists \delta > 0 \text{ tale che } \forall x, y \in A : |x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon.$$

Certamente le funzioni uniformemente continue sono continue. In generale il viceversa non è vero.

Esempio 7. La funzione $f(x) = x^2$ non è uniformemente continua. Infatti, supponiamo per assurdo che ci siano ε, δ che verifichino la definizione di uniforme continuità. Per $x > 0$ prendiamo $y = x + \delta/2$, per cui $|x - y| = \delta/2 < \delta$. Tuttavia si ha

$$|y^2 - x^2| = |x^2 - x^2 - \delta^2/4 - x\delta| = x\delta + \delta^2/4 > x\delta.$$

Allora se si prende $x > \varepsilon/\delta$ si ottiene $|x^2 - y^2| > \varepsilon$.

Lemma 8 (Permanenza del segno). Sia $f : A \rightarrow \mathbb{R}$ una funzione continua, $x_0 \in A$. Se $f(x_0) > l$ (risp. $f(x_0) < l$) allora esiste $r > 0$ tale che $f(x) > l$ (risp. $f(x) < l$) per ogni $x \in [x_0 - r, x_0 + r] \cap A$.

Dimostrazione. Prendiamo $\varepsilon > 0$ tale che $f(x_0) > l + \varepsilon$. Per continuità in x_0 , in corrispondenza di tale $\varepsilon > 0$ possiamo trovare $\delta > 0$ tale che per ogni $x \in A$ risulti

$$|x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon.$$

Se dunque $x \in [x_0 - \delta, x_0 + \delta] \cap A$ risulta $|x - x_0| < \delta$ e quindi $f(x_0) - f(x) < \varepsilon$. Pertanto $f(x) > f(x_0) - \varepsilon > l + \varepsilon - \varepsilon = l$. L'altro caso si dimostra in maniera del tutto analoga. \square

Teorema 9 (degli zeri). Sia $I \subset \mathbb{R}$ un intervallo e $f : I \rightarrow \mathbb{R}$ una funzione continua. Se esistono due punti $a, b \in I$ tali che $f(a) < 0 < f(b)$, allora esiste un punto $x_0 \in I$ tale che $f(x_0) = 0$.

Dimostrazione. Definiamo l'insieme $A := \{x \in [a, b] : f(x) < 0\}$. Osseviamo che $a \in A \neq \emptyset$. Essendo A limitato possiamo considerare $x_0 = \sup A$. Vogliamo ora provare che $f(x_0) = 0$. Infatti, se fosse $f(x_0) < 0$ avremmo $x_0 < b$. Per la permanenza del segno troveremmo $r > 0$ tale che $f(x_0 + r) < 0$. Ma questo contraddice il fatto che x_0 è estremo superiore di A . Se invece fosse $f(x_0) > 0$, per la permanenza del segno troveremmo un $r > 0$ per cui f è strettamente positiva in $[x_0 - r, x_0]$. Per le proprietà dell'estremo superiore possiamo trovare un $x \in A$ tale che $x_0 - r < x \leq x_0$. Ma allora si avrebbe $0 < f(x) < 0$. \square

Corollario 10 (Teorema dei valori intermedi). *Sia $I \subset \mathbb{R}$ un intervallo e $f : I \rightarrow \mathbb{R}$ una funzione continua. Dati due punti $a, b \in I$, la funzione f assume tutti i valori compresi tra $f(a)$ e $f(b)$.*

Dimostrazione. Sia $f(a) < k < f(b)$. Dobbiamo quindi mostrare che esiste un punto $x_0 \in I$ tale che $f(x_0) = k$. Consideriamo la funzione continua $g(x) = f(x) - k$. Allora

$$g(a) = f(a) - k < 0, \quad g(b) = f(b) - k > 0.$$

Per il Teorema degli zeri possiamo concludere che $g(x_0) = 0$ per un opportuno $x_0 \in I$. Ma allora $f(x_0) = k$. \square

Esempio 11. *La funzione $P(x) = 3x^5 - 4x^2 - x + 1$ è continua? Vedremo in seguito che la risposta è affermativa. Osservato allora che*

$$P(0) = 1, \quad P(1) = -1,$$

per il Teorema degli zeri possiamo concludere che il polinomio $P(x)$ ammette almeno una radice.

1.2 Successioni

Definizione 12. *Si dice successione ogni funzione $a : \mathbb{N} \rightarrow \mathbb{R}$. Indicati con $a_0 = a(0), \dots, a_n = a(n)$, si indicherà la successione semplicemente elencandone in maniera compatta i valori e scriveremo*

$$(a_n)_{n \geq 0}.$$

Definizione 13. *Si dice che una successione è monotona crescente (risp. decrescente) se $a_n \leq a_{n+1}$ (risp. $a_{n+1} \leq a_n$) per ogni $n \geq 0$.*

Proposizione 14. *Ogni successione monotona ammette limite.*

Dimostrazione. Supponiamo ad esempio che la successione sia crescente. Sia $l = \sup a_n$. Proviamo che $a_n \rightarrow l$ per $n \rightarrow +\infty$. Infatti, fissato $\varepsilon > 0$, per le proprietà dell'estremo superiore possiamo trovare un indice ν tale che $l - \varepsilon < a_\nu$. Allora, per $n \geq \nu$ si ottiene $|a_n - l| = l - a_n \leq l - a_\nu < \varepsilon$. \square

Esempio 15. $\lim_{n \rightarrow +\infty} \frac{1}{n} = 0$.

Fissiamo $\varepsilon > 0$. Poichè i numeri naturali non sono limitati in \mathbb{R} possiamo trovare $\nu \in \mathbb{N}$ tale che $\nu > \frac{1}{\varepsilon}$. Allora

$$\forall n \geq \nu : \frac{1}{n} \leq \frac{1}{\nu} < \varepsilon.$$

La successione $((-1)^n)_{n \geq 0}$ non è convergente.

Esempio 16. $\lim_{n \rightarrow +\infty} n = +\infty$. *Osserviamo che se due successioni verificano definitivamente*

$$a_n \leq b_n$$

allora se la successione $(a_n)_{n \geq 0}$ è divergente positivamente, anche la $(b_n)_{n \geq 0}$ lo è. Così ad esempio si ha

$$\lim_{n \rightarrow +\infty} n! = +\infty, \quad \lim_{n \rightarrow +\infty} n^n = +\infty.$$

Esercizio

$$\lim_{n \rightarrow +\infty} \sqrt{n}.$$

Esempio 17. $\lim_{n \rightarrow +\infty} \frac{n^2 + 2n - 1}{2n^2 + 1}$.

$$\frac{n^2 + 2n - 1}{2n^2 + 1} = \frac{1 + 2/n - 1/n^2}{2 + 1/n^2} \rightarrow \frac{1}{2}.$$

Esempio 18. $\lim_{n \rightarrow +\infty} \frac{\cos \sqrt{n}}{n} = 0$.

Esercizi

1. $\lim_{n \rightarrow +\infty} \frac{n+2+\cos n}{3-n}$,

2. $\lim_{n \rightarrow +\infty} \frac{n!}{n^n}$,

3. $\lim_{n \rightarrow +\infty} n - \sqrt{n}$,

Definizione 19. Data una successione (a_n) e una successione crescente di interi $k_n \leq k_{n+1}$, la successione (a_{k_n}) si dice *successione estratta* o *sottosuccessione*.

Proposizione 20. Se $a_n \rightarrow l$ allora ogni successione estratta converge allo stesso limite.

Dimostrazione. Fissato $\varepsilon > 0$, definitivamente si ha $|a_n - l| < \varepsilon$. Poichè $k_n \geq n$, definitivamente si ha pure che $|a_{k_n} - l| < \varepsilon$. \square

Definizione 21 (Insiemi compatti). Un insieme $K \subset \mathbb{R}$ si dice *compatto* (per successioni) se ogni successione di elementi di K ammette una estratta che converge ad un elemento di K .

Teorema 22. Ogni intervallo $[a, b] \subset \mathbb{R}$ è compatto.

Dimostrazione. Sia (x_n) una successione di $[a, b]$. Consideriamo l'insieme

$$A = \{x \in [a, b] : x_n < x \text{ per un numero finito di indici } n\}.$$

Chiaramente $a \in A \neq \emptyset$. Poichè A è limitato possiamo considerare $c = \sup A$. Naturalmente $c \leq b$. Facciamo vedere che almeno una successione estratta converge a c . Infatti, se così non fosse, in un piccolo intorno di c si troverebbero soltanto un numero finito di elementi della successione. Ovvero esisterebbe $\delta > 0$ tale che $x_n \in [c - \delta, c + \delta]$ solo per un numero finito di indici n . Ma allora $c + \delta \in A$ e questo contraddice il fatto che $c = \sup A$. \square

Teorema 23 (Bolzano-Weierstrass). Ogni successione limitata ammette una estratta convergente.

Dimostrazione. Dalla limitatezza segue che $x_n \in [-M, M]$ per ogni indice n . Poichè l'intervallo $[-M, M]$ è compatto segue che la successione deve ammettere una estratta convergente. \square

Definizione 24. Un punto x_0 si dice di *accumulazione* per un insieme A se esiste una successione $x_n \in A \setminus \{x_0\}$ tale che $x_n \rightarrow x_0$.

Osserviamo che i punti di accumulazione possono anche non appartenere all'insieme dato. Ad esempio, gli estremi inferiore e superiore di un insieme sono punti di accumulazione. Infatti se $x_0 = \sup A$, $x_0 \notin A$, per $\varepsilon = 1/n > 0$, utilizzando le proprietà dell'estremo superiore possiamo trovare $x_n \in A$ tale che $x_0 - 1/n < x_n \Rightarrow |x_0 - x_n| < 1/n$. Dunque $x_n \rightarrow x_0$. I punti che non sono di accumulazione vengono detti *punti isolati*.

Teorema 25 (di Weierstrass). $f : [a, b] \rightarrow \mathbb{R}$ continua. Allora f è dotata di punti di massimo e di punti di minimo.

Dimostrazione. Sia $m = \inf f$. Allora esiste una successione y_n di $f([a, b])$ tale che $y_n \rightarrow m$. Resta così individuata una successione $x_n \in [a, b]$ tale che $y_n = f(x_n)$. Utilizzando la compattezza possiamo considerare una estratta tale che $x_{k_n} \rightarrow x \in [a, b]$. Poichè la funzione f è continua si ottiene

$$f(x) = \lim_{n \rightarrow +\infty} f(x_{k_n}) = \lim_{n \rightarrow +\infty} y_{k_n} = \lim_{n \rightarrow +\infty} y_n = m.$$

Quindi f è dotata di minimo. In modo analogo si dimostra l'esistenza di un punto di massimo. \square

Teorema 26 (Heine-Cantor). $f : K \rightarrow \mathbb{R}$ continua. Se K è compatto allora la funzione f è uniformemente continua.

Dimostrazione. Supponiamo per assurdo che f non sia uniformemente continua. Allora esisterebbe $\varepsilon > 0$ tale che per ogni $\delta > 0$ si possono trovare punti $x, y \in K$ tali che $|x - y| < \delta$ e $|f(x) - f(y)| \geq \varepsilon$. In corrispondenza di $\delta = 1/n > 0$ possiamo trovare due successioni $x_n, y_n \in K$ tali che $|x_n - y_n| < 1/n$ e $|f(x_n) - f(y_n)| \geq \varepsilon$. Dalla compattezza di K troviamo una estratta $x_{k_n} \rightarrow x_0 \in K$. Allora

$$|y_{k_n} - x_0| \leq |y_{k_n} - x_{k_n}| + |x_{k_n} - x_0| < 1/k_n + |x_{k_n} - x_0|.$$

Pertanto è pure $y_{k_n} \rightarrow x_0$. Essendo f continua le successioni $f(x_{k_n})$ e $f(y_{k_n})$ convergono entrambe a $f(x_0)$. Ma allora la condizione $|f(x_{k_n}) - f(y_{k_n})| \geq \varepsilon$ non può essere soddisfatta. \square

Capitolo 2

Integrazione

Due problemi fondamentali dell'analisi sono la determinazione della tangente ad una curva assegnata e la determinazione dell'area sottesa dalla stessa. Il primo di questi problemi conduce allo studio delle proprietà di differenziabilità delle funzioni mentre il secondo conduce alla teoria dell'integrazione. In questo capitolo ci concentreremo sul problema delle aree. L'idea principale è quella di approssimare l'area mediante delle aree *elementari* e quindi con un procedimento di passaggio al limite.

2.1 Definizione di integrale secondo Riemann

Consideriamo dunque una funzione limitata $f : [a, b] \rightarrow \mathbb{R}$. Un primo passo da compiere è quello di introdurre la nozione di partizione di un intervallo.

Definizione 27 (Partizioni). *Si dice partizione di $[a, b]$ ogni insieme finito*

$$\pi = \{x_0, x_1, \dots, x_n\}$$

tale che $a = x_0 < x_1 < \dots < x_n = b$. Indicheremo con $\mathcal{P}(a, b)$ l'insieme di tutte le partizioni dell'intervallo $[a, b]$.

A questo punto possiamo conderare le aree dei rettangolini associati ad una partizione mediante la seguente

Definizione 28 (Somme inferiori e superiori). *Sia $\pi \in \mathcal{P}(a, b)$. Posti*

$$m_i = \inf\{f(x) : x \in [x_{i-1}, x_i]\}, \quad M_i = \sup\{f(x) : x \in [x_{i-1}, x_i]\},$$

definiamo le somme inferiori e superiori relative alla partizione π rispettivamente mediante:

$$s(f, \pi) = \sum_{i=1}^n m_i(x_i - x_{i-1}), \quad S(f, \pi) = \sum_{i=1}^n M_i(x_i - x_{i-1}).$$

Supponiamo ora che la funzione limitata $f : [a, b] \rightarrow \mathbb{R}$ sia anche positiva e definiamo il sottografico di f come l'insieme

$$\Gamma(f) := \{(x, y) \in \mathbb{R}^2 : x \in [a, b], 0 \leq y \leq f(x)\}.$$

Il problema delle aree da cui siamo partiti consiste dunque nell'assegnare un qualche significato alla parola *area* per $\Gamma(f)$ ed in particolare per figure piane qualsiasi. Quale che sia il significato di quest'area, certamente vorremo che valga la proprietà

$$E \subset F \Rightarrow \text{Area}(E) \leq \text{Area}(F).$$

In particolare, poichè le somme inferiori sono contenute nel sottografico mentre quelle superiori contengono quest'ultimo, avremo

$$s(f, \pi) \leq \text{Area}(\Gamma(f)) \leq S(f, \pi) \quad \forall \pi \in \mathcal{P}(a, b). \quad (2.1)$$

La migliore approssimazione in (2.1) si ottiene considerando l'estremo superiore delle somme inferiori e l'estremo inferiore di quelle superiori. Questo conduce alla seguente

Definizione 29 (Integrale secondo Riemann). *Data $f : [a, b] \rightarrow \mathbb{R}$ limitata si definiscono l'integrale inferiore e superiore della f in $[a, b]$ rispettivamente mediante*

$$s(f) = \sup\{s(f, \pi) : \pi \in \mathcal{P}(a, b)\}, \quad S(f) = \inf\{S(f, \pi) : \pi \in \mathcal{P}(a, b)\}.$$

La funzione f si dice integrabile (secondo Riemann) in $[a, b]$ se $s(f) = S(f)$ e tale valore comune si indica con il simbolo

$$\int_a^b f(x) dx.$$

La Definizione 29 ha un importante significato geometrico. Considerando infatti gli estremi superiori e inferiori nella disuguaglianza (2.1) si ottiene

$$s(f) \leq \text{Area}(\Gamma(f)) \leq S(f).$$

Se dunque f è integrabile si ha che

$$\text{Area}(\Gamma(f)) = \int_a^b f(x) dx.$$

Pertanto, se f è integrabile, sarà naturale definire l'area del sottografico di f mediante il suo integrale.

Esempio 30. *Prendendo la funzione costante $f(x) = c$ su $[a, b]$, dalla definizione di integrale si ricava subito che*

$$\int_a^b f(x) = c(b - a).$$

Infatti, per ogni partizione $\pi = \{x_0, x_1, \dots, x_n\}$, abbiamo che $m_i = c = M_i$. Pertanto

$$s(f, \pi) = \sum_{i=1}^n c(x_i - x_{i-1}) = c(b - a).$$

Allo stesso modo si ottiene $S(f, \pi) = c(b - a)$.

Non tutte le funzioni limitate sono integrabili, come è mostrato dal seguente

Esempio 31. *sia $f : [a, b] \rightarrow \mathbb{R}$ la funzione di Dirichelet definita da*

$$f(x) = \begin{cases} 1 & \text{se } x \in \mathbb{Q}, \\ 0 & \text{altrimenti.} \end{cases}$$

Se $\pi = \{x_0, x_1, \dots, x_n\}$ è una qualsiasi partizione, dato che i numeri razionali e irrazionali sono densi in \mathbb{R} , in qualsiasi intervallino $[\alpha, \beta]$ ci saranno sia numeri razionali che irrazionali. Segue allora che $m_i = 0$, mentre $M_i = 1$. Pertanto si ottiene

$$s(f, \pi) = 0, \quad S(f, \pi) = \sum_{i=1}^n (x_i - x_{i-1}) = (b - a).$$

Dunque

$$s(f) = 0 < b - a = S(f),$$

per cui f non è integrabile (secondo Riemann).

Affinché la nozione di *area* introdotta mediante il processo di integrazione sia significativa occorre trovare una classe abbastanza ampia di funzioni che siano integrabili. A tal fine saranno utili delle proprietà di monotonia delle somme inferiori e superiori. In effetti le somme inferiori e superiori hanno buone proprietà di monotonia:

Lemma 32. *sia $f : [a, b] \rightarrow \mathbb{R}$ limitata e $\pi_1, \pi_2 \in \mathcal{P}(a, b)$. Allora si hanno:*

1. $\pi_1 \subset \pi_2 \Rightarrow s(f, \pi_1) \leq s(f, \pi_2)$,
2. $\pi_1 \subset \pi_2 \Rightarrow S(f, \pi_2) \leq S(f, \pi_1)$,
3. $s(f, \pi_1) \leq S(f, \pi_2)$.

Dimostrazione. Se $\pi_1 \subset \pi_2$ allora π_2 si ottiene da π_1 aggiungendo un numero finito di punti. Supponiamo allora che π_2 abbia esattamente un punto in più di π_1 (il caso generale è del tutto analogo). Detto t tale punto, esso si troverà tra due punti consecutivi $x < y$ di π_1 . Osserviamo che prima di x e dopo di y le due partizioni coincidono. Usando il punto 2 dell'esercizio 3 si ottiene

$$\begin{aligned} s(f, \pi_1) &= s(f, \pi_2) + (y - x) \inf_{[x, y]} f - (t - x) \inf_{[x, t]} f - (y - t) \inf_{[t, y]} f = \\ &= s(f, \pi_2) + (y - t) \inf_{[x, y]} f + (t - x) \inf_{[x, y]} f - (t - x) \inf_{[x, t]} f - (y - t) \inf_{[t, y]} f \leq s(f, \pi_2). \end{aligned}$$

Il punto 2 si dimostra allo stesso modo. Si fissino ora due partizioni π_1, π_2 . Poiché $\pi_1, \pi_2 \subset \pi_1 \cup \pi_2$, dai primi due punti abbiamo

$$s(f, \pi_1) \leq s(f, \pi_1 \cup \pi_2) \leq S(f, \pi_1 \cup \pi_2) \leq s(f, \pi_2).$$

□

Mettiamo in evidenza che dal punto 3 del Lemma 32 segue che

$$s(f) \leq S(f). \quad (2.2)$$

Esempio 33. *Calcoliamo $\int_a^b x dx$. A tal fine dividiamo l'intervallo in n parti uguali considerando la partizione $\pi_n = \{x_0, \dots, x_n\}$ in cui $x_i = a + i \frac{b-a}{n}$. Osserviamo che $m_i = x_{i-1}$ mentre $M_i = x_i$. Tenendo conto della formula della somma dei primi numeri interi, $\sum_{i=0}^n i = \frac{n(n+1)}{2}$, abbiamo che*

$$\begin{aligned} s(f, \pi_n) &= \sum_{i=1}^n x_{i-1} (x_i - x_{i-1}) = \sum_{i=1}^n \left[a + (i-1) \frac{b-a}{n} \right] \left[a + i \frac{b-a}{n} - a - (i-1) \frac{b-a}{n} \right] = \\ &= \sum_{i=1}^n \left[a + (i-1) \frac{b-a}{n} \right] \frac{b-a}{n} = (b-a)a + \frac{(b-a)^2}{n^2} \sum_{i=1}^n (i-1) = (b-a)a + \frac{(b-a)^2}{n^2} \frac{n(n+1)}{2} = \\ &= (b-a)a + \frac{(b-a)^2}{2} \frac{n+1}{n} \xrightarrow{n \rightarrow +\infty} (b-a)a + \frac{(b-a)^2}{2} = \frac{b^2 - a^2}{2}. \end{aligned}$$

Allo stesso modo si ottiene che $\lim_{n \rightarrow +\infty} S(f, \pi_n) = \frac{b^2 - a^2}{2}$. Tenuto conto di (2.2) abbiamo

$$\frac{b^2 - a^2}{2} = \lim_{n \rightarrow +\infty} s(f, \pi_n) \leq s(f) \leq S(f) \leq \lim_{n \rightarrow +\infty} S(f, \pi_n) = \frac{b^2 - a^2}{2}.$$

Pertanto $\int_a^b x dx = s(f) = S(f) = \frac{b^2 - a^2}{2}$.

Naturalmente l'integrale dell'esempio precedente poteva essere calcolato direttamente mediante la geometria elementare. Infatti, posto $f(x) = x$, abbiamo

$$\text{Area}(\Gamma(f)) = \frac{(b-a)^2}{2} + (b-a)a = \frac{b^2 - a^2}{2}.$$

Questo mostra tra le altre cose come possa essere difficoltoso calcolare un integrale usando soltanto le somme superiori ed inferiori anche in casi molto semplici. Avremo pertanto bisogno di strumenti un pò più raffinati. La fatica che faremo sarà però ripagata dalla possibilità di poter integrare funzioni molto più generali.

Enunciamo ora un'importante caratterizzazione delle funzioni integrabili secondo Riemann.

Teorema 34. *Sia $f : [a, b] \rightarrow \mathbb{R}$ una funzione limitata. Le seguenti condizioni sono equivalenti:*

1. f è integrabile secondo Riemann,
2. $\forall \varepsilon > 0 \exists \pi \in \mathcal{P}(a, b)$ t.c. $S(f, \pi) - s(f, \pi) < \varepsilon$.

Dimostrazione. $1 \Rightarrow 2$. Fissiamo $\varepsilon > 0$. Ricordiamo che gli integrali inferiore e superiore sono definiti rispettivamente da un'estremo superiore e da un'estremo inferiore. Allora, usando le caratterizzazioni date dall'esercizio 1 e dall'esercizio 2, possiamo trovare due partizioni $\pi_1, \pi_2 \in \mathcal{P}(a, b)$ tali che

$$s(f) < s(f, \pi_1) + \frac{\varepsilon}{2}, \quad S(f, \pi_2) \leq S(f) + \frac{\varepsilon}{2}.$$

Presa allora una qualsiasi partizione π tale che $\pi_1, \pi_2 \subset \pi$, ad esempio $\pi = \pi_1 \cup \pi_2$, tenuto conto del Lemma 32 e del fatto che $s(f) = S(f)$ si ottiene

$$S(f, \pi) - s(f, \pi) \leq S(f, \pi_2) - s(f, \pi_1) < S(f) + \frac{\varepsilon}{2} - s(f) + \frac{\varepsilon}{2} = \varepsilon.$$

Proviamo adesso che $2 \Rightarrow 1$. Fissiamo ancora $\varepsilon > 0$. Dalla condizione 2 troviamo una $\pi \in \mathcal{P}(a, b)$ tale che

$$0 \leq S(f) - s(f) \leq S(f, \pi) - s(f, \pi) < \varepsilon.$$

Dall'arbitrarietà di $\varepsilon > 0$ si ottiene che $S(f) = s(f)$. □

Ora siamo in grado di individuare alcune classi di funzioni integrabili; una di queste è costituita dalle funzioni monotone.

Proposizione 35. *Ogni funzione $f : [a, b] \rightarrow \mathbb{R}$ monotona è integrabile.*

Dimostrazione. Prendiamo la partizione $\pi_n = \{x_0, \dots, x_n\}$ dell'esempio 33 in cui $x_i = a + i \frac{b-a}{n}$. Supponiamo per esempio che la f sia crescente. Allora abbiamo che

$$m_i = f(x_{i-1}), \quad M_i = f(x_i).$$

Pertanto possiamo valutare

$$\begin{aligned} s(f, \pi_n) &= \sum_{i=1}^n f(x_{i-1})(x_i - x_{i-1}) = \sum_{i=1}^n f(x_{i-1}) \left[a + i \frac{b-a}{n} - \left(a + (i-1) \frac{b-a}{n} \right) \right] = \\ &= \frac{b-a}{n} \sum_{i=1}^n f(x_{i-1}) = \frac{b-a}{n} \sum_{i=0}^{n-1} f(x_i) = \frac{b-a}{n} \left(f(a) + \sum_{i=1}^{n-1} f(x_i) \right), \\ S(f, \pi_n) &= \frac{b-a}{n} \sum_{i=1}^n f(x_i) = \frac{b-a}{n} \left(f(b) + \sum_{i=1}^{n-1} f(x_i) \right). \end{aligned}$$

Dunque

$$S(f, \pi_n) - s(f, \pi_n) = \frac{b-a}{n} (f(b) - f(a)) \xrightarrow{n \rightarrow +\infty} 0.$$

L'integrabilità segue allora dal Teorema 34. □

Un'altra classe importante di funzioni integrabili è costituita dalle funzioni continue. Ricordiamo che se $I \subset \mathbb{R}$ una funzione $f : I \rightarrow \mathbb{R}$ è continua in I se verifica la condizione

$$\forall x \in I, \lim_{y \rightarrow x} f(y) = f(x). \quad (2.3)$$

In termini di (ε, δ) la condizione (2.3) è equivalente a chiedere che

$$\forall x \in I, \forall \varepsilon > 0 \exists \delta > 0 \text{ t.c. } \forall y \in I, |x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon. \quad (2.4)$$

Osserviamo che il numero $\delta > 0$ nella condizione (2.4) in generale può dipendere non solo dalla scelta di $\varepsilon > 0$ ma anche dal punto $x \in I$. Se si può scegliere un $\delta > 0$ indipendentemente dal punto $x \in I$ allora si dice che la funzione f è uniformemente continua in I . Dunque f è uniformemente continua se verifica:

$$\forall \varepsilon > 0 \exists \delta > 0 \text{ t.c. } \forall x, y \in I, |x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon. \quad (2.5)$$

Esercizio 36. *Tutte le funzioni Lipschitziane sono uniformemente continue. Le funzioni di classe $C^1(I)$ con I chiuso e limitato sono Lipschitziane e quindi uniformemente continue.*

Il Teorema di Heine-Cantor (vedi per esempio il Teorema 6.41 in [?]) stabilisce che le funzioni continue definite su intervalli chiusi e limitati sono in realtà uniformemente continue.

Il teorema seguente identifica un'altra classe importante di funzioni continue.

Teorema 37. *Ogni funzione $f : [a, b] \rightarrow \mathbb{R}$ continua è integrabile.*

Dimostrazione. L'intervallo $[a, b]$ è chiuso e limitato e per il Teorema di Weierstrass f risulta essere limitata e come appena ricordato uniformemente continua. Fissato $\varepsilon > 0$ scegliamo un $\delta > 0$ che soddisfi la condizione (2.5). Prendiamo una partizione $\pi = \{x_0, \dots, x_n\}$ abbastanza fitta di modo che $|x_i - x_{i-1}| < \delta$. Ancora per il Teorema di Weierstrass, in ogni intervallo $[x_i, x_{i-1}]$ la funzione f ammette minimo m_i e massimo M_i e questi valori sono raggiunti in alcuni punti, diciamo rispettivamente y_i e z_i . Chiaramente $|y_i - z_i| \leq x_i - x_{i-1} < \delta$. Dalla (2.5) otteniamo

$$M_i - m_i = f(y_i) - f(z_i) < \varepsilon.$$

Dunque possiamo valutare che

$$\begin{aligned} S(f, \pi) - s(f, \pi) &= \sum_{i=1}^n M_i(x_i - x_{i-1}) - \sum_{i=1}^n m_i(x_i - x_{i-1}) = \\ &= \sum_{i=1}^n (M_i - m_i)(x_i - x_{i-1}) < \varepsilon \sum_{i=1}^n (x_i - x_{i-1}) = \varepsilon(b - a). \end{aligned}$$

L'integrabilità segue allora dal Teorema 34. □

Una proprietà elementare ma importante delle funzione integrabili è il fatto che l'integrazione è una operazione lineare.

Teorema 38. *Siano $f, g : [a, b] \rightarrow \mathbb{R}$ due funzioni integrabili. Allora si hanno:*

1. $f + g$ è integrabile e

$$\int_a^b (f(x) + g(x))dx = \int_a^b f(x)dx + \int_a^b g(x)dx,$$

2. se $\lambda \in \mathbb{R}$ allora λf è integrabile e

$$\int_a^b \lambda f(x)dx = \lambda \int_a^b f(x)dx.$$

Dimostrazione. A scopo illustrativo proviamo il punto 1. Fissiamo una qualsiasi partizione $\pi = \{x_0, \dots, x_n\} \in \mathcal{P}(a, b)$. Dal punto 2 dell'esercizio 3 abbiamo che

$$\begin{aligned} s(f, \pi) + s(g, \pi) &= \sum_{i=1}^n (x_i - x_{i-1}) \left(\inf_{[x_{i-1}, x_i]} f + \inf_{[x_{i-1}, x_i]} g \right) \leq \\ &\leq \sum_{i=1}^n (x_i - x_{i-1}) \left(\inf_{[x_{i-1}, x_i]} (f + g) \right) = s(f + g, \pi). \end{aligned} \quad (2.6)$$

Similmente, usando il punto 1 dell'esercizio 3 si ottiene

$$S(f + g, \pi) \leq S(f, \pi) + S(g, \pi). \quad (2.7)$$

Dunque, se π_1, π_2 sono due qualsiasi partizioni di $[a, b]$, grazie alla Proposizione 32 e alla (2.6) abbiamo

$$s(f, \pi_1) + s(g, \pi_2) \leq s(f, \pi_1 \cup \pi_2) + s(g, \pi_1 \cup \pi_2) \leq s(f + g, \pi_1 \cup \pi_2) \leq s(f + g).$$

Passando all'estremo superiore rispetto a π_1 e π_2 si ottiene

$$\int_a^b f(x)dx + \int_a^b g(x)dx \leq s(f + g).$$

Analogamente, usando la (2.7) si ottiene

$$S(f + g) \leq \int_a^b f(x)dx + \int_a^b g(x)dx.$$

In definitiva abbiamo ottenuto che

$$S(f + g) \leq \int_a^b f(x)dx + \int_a^b g(x)dx \leq s(f + g),$$

e poichè $s(f + g) \leq S(f + g)$ allora $f + g$ è integrabile e vale l'uguaglianza

$$\int_a^b (f(x) + g(x))dx = \int_a^b f(x)dx + \int_a^b g(x)dx.$$

Se $\lambda \geq 0$ per il punto 5 dell'esercizio 3 abbiamo

$$s(\lambda f, \pi) = \lambda s(f, \pi), \quad S(\lambda f, \pi) = \lambda S(f, \pi).$$

Passando rispettivamente all'estremo superiore e all'estremo inferiore si ottiene

$$s(\lambda f) = \lambda s(f) = \lambda \int_a^b f(x)dx, \quad S(\lambda f) = \lambda S(f) = \lambda \int_a^b f(x)dx,$$

da cui segue l'asserto. Se invece $\lambda < 0$, usando il punto 6 dell'esercizio 3 otteniamo

$$s(\lambda f, \pi) = \lambda S(f, \pi), \quad S(\lambda f, \pi) = \lambda s(f, \pi).$$

Allora

$$s(\lambda f) = \sup_{\pi \in \mathcal{P}(a, b)} s(\lambda f, \pi) = \sup_{\pi \in \mathcal{P}(a, b)} \lambda S(f, \pi) = \lambda \inf_{\pi \in \mathcal{P}(a, b)} S(f, \pi) = \lambda S(f) = \lambda \int_a^b f(x)dx,$$

$$S(\lambda f) = \inf_{\pi \in \mathcal{P}(a, b)} S(\lambda f, \pi) = \inf_{\pi \in \mathcal{P}(a, b)} \lambda s(f, \pi) = \lambda \sup_{\pi \in \mathcal{P}(a, b)} s(f, \pi) = \lambda s(f) = \lambda \int_a^b f(x)dx.$$

□

2.2 Cenni alla misura e all'integrale di Lebesgue

Nella sezione precedente abbiamo definito l'area del sottografico di una funzione limitata approssimando quest'ultimo con rettangolini *inscritti* e *circoscritti*. Questa operazione può naturalmente essere eseguita per una qualsiasi figura piana e produce una misura nel piano, e più precisamente la misura di Peano-Jordan in \mathbb{R}^2 . Dunque diremo che un insieme limitato $A \subset \mathbb{R}^2$ è misurabile secondo Peano-Jordan se l'approssimazione con i rettangolini *inscritti* coincide con l'approssimazione mediante i rettangolini *circoscritti* e tale valore comune è la misura di A che indicheremo con $\mu(A)$. Pertanto il concetto di *area* introdotto nella sezione precedente si può formulare dicendo che se $f : [a, b] \rightarrow \mathbb{R}$ è integrabile secondo Riemann allora

$$\int_a^b f(x)dx = \mu(\Gamma(f)).$$

Una misura è in generale una funzione $\mu : \mathcal{M} \rightarrow [0, +\infty]$ definita su una famiglia di insiemi misurabili \mathcal{M} che verifica:

1. $\emptyset \in \mathcal{M}$,
2. $A \in \mathcal{M} \Rightarrow A^c \in \mathcal{M}$,
3. $A_n \in \mathcal{M} \Rightarrow \bigcup_{n=0}^{+\infty} A_n \in \mathcal{M}$.

Le principali proprietà di una misura sono:

1. $\mu(\emptyset) = 0$,
2. (additività numerabile) $\mu\left(\bigcup_{n=0}^{+\infty} A_n\right) = \sum_{n=0}^{+\infty} \mu(A_n)$, se gli A_n sono a due a due disgiunti,
3. (monotonia) $A \subset B \Rightarrow \mu(A) \leq \mu(B)$.

Poichè $\mu(\{a\}) = 0$ (perchè?) deduciamo che

$$\int_a^a f(x) = \mu(\{a\}) = 0.$$

La proprietà di additività 2 delle misure si traduce nella seguente

Proposizione 39 (Spezzamento). *Sia $f : [a, b] \rightarrow \mathbb{R}$ integrabile. Allora per ogni $c \in [a, b]$ si ha*

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx.$$

In particolare l'additività ci dice che il valore dell'integrale non cambia rispetto ad insiemi di misura nulla e quindi ai valori che la funzione assume su questi insiemi. In particolare, possiamo ad esempio trascurare una qualsiasi infinità numerabile di punti dell'intervallo $[a, b]$. Pertanto le funzioni che sono continue fatta eccezione al più per una infinità numerabile di punti sono integrabili. Questo è il caso ad esempio delle funzioni monotone (si veda per esempio la Proposizione A6.13 in [1]). La misura di Peano-Jordan non è l'unico modo per misurare le aree. Anzi, da certi punti di vista, pur avendo il pregio della semplicità, la teoria dell'integrazione secondo Riemann non è del tutto soddisfacente. Per diversi motivi, a partire dalla fine del diciannovesimo secolo, i matematici cominciarono a cercare una nozione di integrale più generale che consentisse, tra le altre cose, di integrare funzioni molto discontinue come ad esempio la funzione di Dirichelet. La nozione che ottenne maggior successo e oggi più diffusa è quella dovuta al matematico francese Henri Lebesgue a partire dalla sua tesi di dottorato del 1902. In effetti, le funzioni integrabili secondo Riemann sono integrabili secondo Lebesgue e sono *praticamente* funzioni continue. Sussiste infatti il seguente

Teorema 40 (Lebesgue-Vitali). *Una funzione $f : [a, b] \rightarrow \mathbb{R}$ è integrabile secondo Riemann se e soltanto se l'insieme dei suoi punti di discontinuità ha misura nulla secondo Lebesgue.*

Per i nostri scopi sarà senz'altro sufficiente la nozione di integrale secondo Riemann. Dunque da ora in poi potremo senza troppi problemi restringerci a considerare funzioni continue dove le due nozioni di integrale coincidono. Per maggiori dettagli sulla teoria della misura e dell'integrazione si può consultare [?].

Terminiamo questa sezione facendo alcune ossevazioni semplici ma importanti. Intanto osserviamo che la proprietà di monotonia della misura si traduce nella seguente

Proposizione 41. *Siano $f, g : [a, b] \rightarrow \mathbb{R}$ due funzioni integrabili. Allora*

$$\forall x \in [a, b], \quad f(x) \leq g(x) \Rightarrow \int_a^b f(x)dx \leq \int_a^b g(x)dx.$$

Esercizio 42. *Si provi che per ogni funzione continua $f : [a, b] \rightarrow \mathbb{R}$ si ha:*

$$\left| \int_a^b f(x)dx \right| \leq \int_a^b |f(x)|dx.$$

Un'altra proprietà importante è espressa dal seguente

Teorema 43 (della media integrale). *Sia $f : [a, b] \rightarrow \mathbb{R}$ una funzione integrabile. Allora si ha*

$$\inf_{[a,b]} f \leq \frac{1}{b-a} \int_a^b f(x)dx \leq \sup_{[a,b]} f.$$

Se inoltre la funzione f è continua, allora esiste $z \in [a, b]$ tale che

$$f(z) = \frac{1}{b-a} \int_a^b f(x)dx.$$

Dimostrazione. Per ogni $x \in [a, b]$ si ha

$$\inf_{[a,b]} f \leq f(x) \leq \sup_{[a,b]} f.$$

Integrando otteniamo

$$(b-a) \inf_{[a,b]} f \leq \int_a^b f(x)dx \leq (b-a) \sup_{[a,b]} f,$$

da cui si ottiene la tesi dividendo per $(b-a)$. Nel caso in cui la funzione f sia continua, a causa del Teorema dei valori intermedi, la funzione assume tutti i valori compresi tra il suo minimo ed il suo massimo. Pertanto in particolare assumerà il valore

$$\frac{1}{b-a} \int_a^b f(x)dx.$$

□

Esercizio 44. *Si provi che per ogni funzione continua $f : [a, b] \rightarrow \mathbb{R}$ e per ogni $x_0 \in [a, b]$ si ha:*

$$\lim_{x \rightarrow x_0} \frac{1}{x-x_0} \int_{x_0}^x f(t)dt = f(x_0).$$

2.3 Il Teorema fondamentale del calcolo integrale

In questa sezione ci occuperemo di mettere in relazione l'operazione di integrazione con quella di derivazione. Questi due concetti apparentemente slegati sono in realtà intimamente connessi e il legame tra i due è costituito appunto dal Teorema fondamentale del calcolo integrale. Come già osservato, calcolare un integrale ricorrendo soltanto alla definizione è in genere una via impraticabile. Il Teorema fondamentale del calcolo è inoltre basilare perchè introduce degli strumenti teorici che permettono di calcolare molti tipi di integrali.

In particolare ci porremo due domande principali:

1. Cosa succede se deriviamo un integrale?
2. Cosa succede se integriamo una derivata?

Tra le altre cose saremo poi condotti a riconoscere che integrazione e derivazione sono due operazioni in un certo senso una inversa dell'altra. Per prima cosa avremo bisogno di definire l'integrazione su intervalli orientati.

Definizione 45. Se $b < a$ e f è integrabile in $[b, a]$ definiamo:

$$\int_a^b f(x)dx = - \int_b^a f(x)dx.$$

Esercizio 46. Si verifichi che il Teorema della media, la proprietà di linearità e di spezzamento continuano a valere indipendentemente dall'ordine degli estremi di integrazione.

Teorema 47 (fondamentale del calcolo integrale). Sia I un intervallo e $f : I \rightarrow \mathbb{R}$ una funzione continua. Fissato $a \in I$ si consideri la funzione $F : I \rightarrow \mathbb{R}$ definita da

$$F(x) = \int_a^x f(t)dt.$$

Allora la F è di classe C^1 e per ogni $x \in I$ si ha $F'(x) = f(x)$.

Dimostrazione. Fissiamo un punto $x_0 \in I$. Per la proprietà di spezzamento abbiamo che

$$F(x) - F(x_0) = \int_a^x f(t) - \int_a^{x_0} f(t) = \int_a^x f(t) + \int_{x_0}^a f(t) = \int_{x_0}^x f(t),$$

da cui, se $x \neq x_0$, otteniamo

$$\frac{F(x) - F(x_0)}{x - x_0} = \frac{1}{x - x_0} \int_{x_0}^x f(t).$$

Passando al limite per $x \rightarrow x_0$ (esercizio 44) si ha che la F è derivabile e $F'(x_0) = f(x_0)$. □

Il Teorema fondamentale del calcolo integrale dunque ha dato risposta alla nostra prima domanda affermando che

$$\frac{d}{dx} \int_a^x f(t)dt = f(x).$$

Esercizio 48. Sapreste calcolare $\frac{d}{dx} \int_{\alpha(x)}^{\beta(x)} f(t)dt$ per α, β due qualsiasi funzioni derivabili? Quale esempio si calcoli la derivata della funzione $\int_x^{x^2} e^{-t^2} dt$.

Il Teorema fondamentale del calcolo inoltre asserisce che ogni funzione continua è dotata di primitiva, secondo la seguente

Definizione 49 (Primitive). *Si dice che una funzione F è una primitiva di f se F è derivabile e $F' = f$.*

Osserviamo che se una funzione ammette una primitiva F allora ne ammette infinite altre poichè la funzione $F + c$ è ancora una primitiva per ogni costante $c \in \mathbb{R}$. Si indica con il simbolo $\int f(x)dx$ l'insieme di tutte le primitive della f . Tale simbologia, nonchè l'utilità della nozione di primitiva è giustificata dal seguente

Teorema 50 (di Torricelli). *Sia f una funzione continua su di un intervallo I . Allora tutte le primitive di f differiscono per una costante. Inoltre, se G è una primitiva di f allora per ogni $a, b \in I$ si ha*

$$\int_a^b f(x)dx = G(b) - G(a).$$

Dimostrazione. Siano F, G due primitive di f . Allora $(F - G)' = F' - G' = f - f = 0$. D'altra parte, per ogni $x, y \in I$ il Teorema di Lagrange implica che

$$(F(x) - G(x)) - (F(y) - G(y)) = (F - G)'(\zeta)(x - y) = 0,$$

per cui $F - G$ è costante. Per dimostare il secondo punto prendiamo la funzione

$$F(x) = \int_c^x f(t)dt$$

che sappiamo essere una primitiva di f . Allora

$$\begin{aligned} \int_a^b f(t)dt &= \int_a^c f(t)dt + \int_c^b f(t)dt = \int_c^b f(t)dt - \int_c^a f(t)dt = \\ &= F(b) - F(a) = (G(b) + c) - (G(a) + c) = G(b) - G(a). \end{aligned}$$

□

Il Teorema di Torricelli risponde alla nostra seconda domanda poichè si riconosce immediatamente che se f è di classe C^1 allora

$$\int_a^b f'(x)dx = f(b) - f(a). \quad (2.8)$$

L'interesse del Teorema di Torricelli è poi evidente giacché riconduce il calcolo degli integrali al problema di determinare una primitiva della funzione assegnata. Nel seguito useremo la comoda notazione $[F(x)]_a^b = F(b) - F(a)$.

Esempio 51. *Calcoliamo $\int_a^b x dx$. Poichè una primitiva di x è ad esempio $\frac{x^2}{2}$ si ha*

$$\int_a^b x dx = \left[\frac{x^2}{2} \right]_a^b = \frac{b^2 - a^2}{2}$$

come avevamo già faticosamente accertato.

$$\int_0^{\frac{\pi}{4}} \frac{1}{\cos^2 x} dx = [\tan x]_0^{\frac{\pi}{4}} = \tan \frac{\pi}{4} - \tan 0 = 1.$$

$$\int_1^2 \frac{\log x}{x} dx = \left[\frac{1}{2} \log^2 x \right]_1^2 = \frac{1}{2} \log^2 2 - \frac{1}{2} \log^2 1 = \frac{1}{2} \log^2 2.$$

$$\int_{\frac{\pi}{2}}^1 \frac{\cos x}{\sin x} dx = [\log |\sin x|]_{\frac{\pi}{2}}^1 = \log |\sin 1| - \log \left| \sin \left(\frac{\pi}{2} \right) \right| = \log |\sin 1|.$$

$$\int_0^1 \frac{1}{\cosh x} dx = \int_0^1 \frac{2}{e^x + e^{-x}} dx = 2 \int_0^1 \frac{e^x}{1 + (e^x)^2} dx = 2 [\arctan e^x]_0^1 = 2(\arctan e - \arctan 1).$$

2.4 Alcuni metodi di integrazione

Gli esempi precedenti mostrano come possa essere, almeno in linea di principio, calcolati una vasta classe di integrali. Tuttavia non sempre è banale scrivere la funzione integranda come derivata di qualcosa e spesso occorrono intuizione e artifici. In effetti, mentre la conoscenza delle derivate elementari consente di derivare funzioni anche complicatissime senza sostanziali difficoltà, per l'integrazione non esistono tecniche di routine che vadano bene in tutti i casi e pertanto la tecnica dell'integrazione richiede esperienza ed intuizione. In questa sezione svilupperemo alcuni risultati utili per il calcolo esplicito di primitive.

2.4.1 Integrazione per parti

Teorema 52 (Formula di integrazione per parti). *Siano $f, g : I \rightarrow \mathbb{R}$ due funzioni di classe \mathcal{C}^1 . Allora per ogni $a, b \in I$ si ha*

$$\int_a^b f'(x)g(x)dx = [f(x)g(x)]_a^b - \int_a^b f(x)g'(x)dx. \quad (2.9)$$

Dimostrazione. Poniamo $h = fg$. Per il Teorema di Torricelli abbiamo che

$$\int_a^b h'(x)dx = [h(x)]_a^b.$$

Per la regola di derivazione del prodotto otteniamo che

$$\int_a^b h'(x)dx = \int_a^b f'(x)g(x)dx + \int_a^b f(x)g'(x)dx.$$

Combinando le due uguaglianze segue il risultato. \square

La formula di integrazione per parti è molto utile in quanto permette di *scaricare* l'operazione di derivazione da una funzione all'altra. Una larga parte dell'analisi matematica moderna ha a che fare con generalizzazioni di questa formula. Vediamo ora qualche esempio del suo impiego.

$$\int_0^1 xe^x dx = \int_0^1 D(e^x)x dx = [xe^x]_0^1 - \int_0^1 e^x dx = e - [e^x]_0^1 = e - (e - 1) = 1.$$

Osserviamo che avremmo potuto benissimo scrivere

$$\int_0^1 xe^x dx = \int_0^1 D\left(\frac{1}{2}x^2\right)e^x dx = \left[\frac{1}{2}x^2e^x\right]_0^1 - \int_0^1 \frac{1}{2}x^2e^x dx,$$

ma questa volta l'operazione di integrazione si è ulteriormente complicata rispetto alla situazione di partenza. Dunque a volte può anche essere utile percorrere strade diverse in modo da trovare la strategia vincente. Naturalmente la relativa formula per l'integrale indefinito è data da

$$\int f'(x)g(x)dx = f(x)g(x) - \int f(x)g'(x)dx + c.$$

Iterando l'esempio precedente si può per esempio calcolare $\int x^n e^x dx$.

Esempio 53. *Calcoliamo $\int \log x dx$. Poichè $D(x) = 1$ integrando per parti otteniamo*

$$\int \log x dx = \int D(x) \log x dx = x \log x - \int x \frac{1}{x} dx = x \log x - x + c.$$

Ripetendo la procedura più volte si può calcolare $\int \log^n x dx$.

Esempio 54. Ancora integrando per parti calcoliamo

$$\begin{aligned}\int x^n \log x dx &= \frac{x^{n+1}}{n+1} \log x - \int \frac{x^{n+1}}{n+1} \frac{1}{x} dx = \frac{x^{n+1}}{n+1} \log x - \frac{1}{n+1} \int x^n dx = \\ &= \frac{x^{n+1}}{n+1} \log x - \frac{1}{(n+1)^2} x^{n+1} dx + c = \frac{x^{n+1}}{n+1} \left(\log x - \frac{1}{n+1} \right) + c.\end{aligned}$$

A volte può risultare utile sviluppare una procedura ciclica in modo da far ricomparire lo stesso integrale di partenza con segno opposto, come mostra il seguente

Esempio 55.

$$\int \sin^2 x dx = \int \sin x \sin x dx = -\cos x \sin x + \int \cos^2 x dx = -\cos x \sin x + \int (1 - \sin^2 x) dx,$$

da cui otteniamo che

$$\int \sin^2 x dx = -\cos x \sin x + x - \int \sin^2 x dx.$$

Portando l'integrale del secondo membro al primo membro dell'equazione si ottiene

$$\int \sin^2 x dx = \frac{1}{2}(x - \cos x \sin x) + c.$$

In modo analogo si possono calcolare $\int \sin^n x dx$ e $\int \cos^n x dx$.

Esempio 56.

$$\begin{aligned}\int \arctan x dx &= \int D(x) \arctan x dx = x \arctan x - \int \frac{x}{1+x^2} dx = \\ &= x \arctan x - \frac{1}{2} \int \frac{2x}{1+x^2} dx = x \arctan x - \frac{1}{2} \log(1+x^2) + c.\end{aligned}$$

Esempio 57. Calcoliamo l'area del semicerchio di centro l'origine e raggio 1. Poichè l'equazione corrispondente è $x^2 + y^2 = 1$, si tratta di calcolare:

$$\begin{aligned}\int_{-1}^1 \sqrt{1-x^2} dx &= \left[x\sqrt{1-x^2} \right]_{-1}^1 + \int_{-1}^1 \frac{x^2}{\sqrt{1-x^2}} dx = \int_{-1}^1 \frac{x^2+1-1}{\sqrt{1-x^2}} dx = \\ &= \int_{-1}^1 \frac{1}{\sqrt{1-x^2}} dx - \int_{-1}^1 \frac{1-x^2}{\sqrt{1-x^2}} dx = [\arcsin x]_{-1}^1 - \int_{-1}^1 \sqrt{1-x^2} dx.\end{aligned}$$

Pertanto

$$\int_{-1}^1 \sqrt{1-x^2} dx = \frac{1}{2}(\arcsin 1 - \arcsin(-1)) = \frac{\pi}{2}.$$

2.4.2 Integrazione per sostituzione

Teorema 58 (Formula di cambiamento delle variabili). Sia $f : I \rightarrow \mathbb{R}$ continua. Siano inoltre J un altro intervallo e $\varphi : J \rightarrow I$ una funzione di classe \mathcal{C}^1 . Allora per ogni $a, b \in I$ si ha

$$\int_a^b f(x) dx = \int_{\varphi^{-1}(a)}^{\varphi^{-1}(b)} f(\varphi(t)) \varphi'(t) dt.$$

Dimostrazione. Dal Teorema fondamentale del calcolo (vedi esercizio 48) abbiamo che

$$\frac{d}{dt} \int_a^{\varphi(t)} f(x) dx = f(\varphi(t)) \varphi'(t).$$

Allora si ottiene che

$$\int_u^v f(\varphi(t)) \varphi'(t) dt = \left[\int_a^{\varphi(t)} f(x) dx \right]_u^v = \int_a^{\varphi(v)} f(x) dx - \int_a^{\varphi(u)} f(x) dx = \int_{\varphi(u)}^{\varphi(v)} f(x) dx.$$

Nel caso in cui φ sia invertibile abbiamo che

$$\int_a^b f(x) dx = \int_{\varphi(\varphi^{-1}(a))}^{\varphi(\varphi^{-1}(b))} f(x) dx = \int_{\varphi^{-1}(a)}^{\varphi^{-1}(b)} f(\varphi(t)) \varphi'(t) dt.$$

Se φ non è invertibile esisteranno più punti, diciamo u, v, z, w tali che $a = \varphi(u) = \varphi(w)$ e $b = \varphi(v) = \varphi(z)$. Dobbiamo quindi dimostrare che la formula appena dimostrata non dipende dalla scelta dei punti $u, v \in J$. Infatti

$$\begin{aligned} \int_u^v f(\varphi(t)) \varphi'(t) dt &= \left[\int_a^{\varphi(x)} f(t) dt \right]_u^v = \int_{\varphi(u)}^{\varphi(v)} f(t) dt = \int_a^b f(x) dx = \\ &= \left[\int_a^{\varphi(x)} f(t) dt \right]_w^z = \int_w^z f(\varphi(t)) \varphi'(t) dt. \end{aligned}$$

□

La formula precedente è nota anche come metodo di sostituzione in quanto ci dice che se sostituiamo la variabile x mediante l'espressione $x = \varphi(t)$, allora l'integrale si trasforma mediante il fattore differenziale di cambiamento $\varphi'(t)$ e modificando opportunamente gli estremi di integrazione.

Esempio 59. Calcoliamo $\int_0^1 \frac{x}{\sqrt{1+x}} dx$. Poniamo $x = \varphi(t) = t^2 - 1$. Allora $\varphi^{-1}(0) = 1$, mentre $\varphi^{-1}(1) = \sqrt{2}$. Dalla formula di cambiamento di variabili otteniamo

$$\int_0^1 \frac{x}{\sqrt{1+x}} dx = \int_1^{\sqrt{2}} \frac{t^2 - 1}{t} 2t dt = \int_1^{\sqrt{2}} 2(t^2 - 1) dt = 2 \left[\frac{t^3}{3} - t \right]_1^{\sqrt{2}} = 2 \left(\frac{\sqrt{8}}{3} - \sqrt{2} - \frac{1}{3} + 1 \right).$$

L'efficacia della sostituzione nell'esempio precedente è consistita nell'eliminazione della radice dal denominatore della funzione integranda. Da questo punto di vista è più naturale porre $x + 1 = t^2$, che però non è della forma $x = \varphi(t)$. Allora sarà conveniente considerare trasformazioni di classe \mathcal{C}^1 del tipo

$$p(x) = q(t), \tag{2.10}$$

a patto di considerare funzioni p invertibili con derivata sempre diversa da zero. Infatti in tal caso la relazione (2.10) è equivalente a porre

$$x = p^{-1}(q(t)) := \varphi(t).$$

Osserviamo che in tal caso dalla formula di derivazione della funzione inversa il fattore differenziale diventa

$$\varphi'(t) = \frac{q'(t)}{p'(p^{-1}(q(t)))} = \frac{q'(t)}{p'(x)}.$$

Un utile strategia per ricordare quanto sopra (ma solo in maniera formale!) è di dedurre da (2.10) 'differenziando'

$$p'(x) dx = q'(t) dt,$$

e quindi ricordarsi che il ‘termine’ dx va sostituito con $\frac{q'(t)}{p'(x)}dt$. La versione per l’integrale indefinito della formula di cambiamento di variabili è data da

$$\int f(x)dx \underset{x=\varphi(t)}{=} \int f(\varphi(t))\varphi'(t)dt,$$

in modo da ricordare che alla fine la primitiva va calcolata nei punti $x = \varphi(t)$. Tale notazione è utile anche negli integrali indefiniti specie quando si fanno più sostituzioni.

Esempio 60. Calcoliamo $\int e^x \cos(e^x)dx$. Poniamo $e^x = t$. Allora

$$\int e^x \cos(e^x)dx \underset{e^x=t}{=} \int t \cos t \frac{1}{t} dt \underset{e^x=t}{=} \sin t + c = \sin(e^x) + c.$$

Esempio 61.

$$\int_{e^{\frac{\pi}{4}}}^{e^{\frac{\pi}{2}}} \frac{\sin^2(\log x) \cos(\log x)}{x} dx \underset{x=e^t}{=} \int_{\frac{\pi}{4}}^{\frac{\pi}{2}} \sin^2 t \cos t dt \underset{\sin t=s}{=} \int_{\frac{\sqrt{2}}{2}}^1 s^2 ds = \left[\frac{s^3}{3} \right]_{\frac{\sqrt{2}}{2}}^1 = \frac{1}{3} - \frac{\sqrt{8}}{6}.$$

2.4.3 Integrazione di funzioni razionali

In questa sezione vogliamo sviluppare una strategia per calcolare integrali in cui la funzione integranda sia il rapporto tra due polinomi. Siano dunque

$$A(x) = a_0x^m + a_1x^{m-1} + \dots + a_m,$$

$$B(x) = b_0x^n + b_1x^{n-1} + \dots + b_n,$$

due polinomi a coefficienti reali e consideriamo la funzione razionale

$$f(x) = \frac{A(x)}{B(x)}.$$

Naturalmente possiamo supporre che i polinomi siano primi tra loro e che $m < n$ poiché altrimenti basterebbe effettuare la divisione. Evidentemente, la difficoltà principale in questi tipi di funzioni integrande è la presenza del denominatore. L’idea principale è quella di scrivere il rapporto $\frac{A(x)}{B(x)}$ in una forma vantaggiosa. Per fare questo articoliamo una procedura in vari passi.

Passo 1. Calcoliamo le n radici di $B(x)$. Queste ultime possono essere sia reali che complesse e vanno contate con la rispettiva molteplicità. In particolare se $z = \alpha + i\beta$ è una radice complessa di B è noto (e si verifica facilmente) che anche il coniugato $\bar{z} = \alpha - i\beta$ è una radice di B con la stessa molteplicità. Inoltre si ha che

$$(x - z)(x - \bar{z}) = (x - (\alpha + i\beta))(x - (\alpha - i\beta)) = (x - \alpha)^2 + \beta^2.$$

Indichiamo con x_j ($j = 1, \dots, r$) le radici reali di B e con k_j le rispettive molteplicità, mentre con z_j ($j = 1, \dots, s$) e h_j indichiamo le radici complesse e la rispettiva molteplicità. Pertanto avremo che $k_1 + \dots + k_r + 2h_1 + \dots + 2h_s = n$.

Passo 2. Supponiamo che $b_0 = 1$, come si può sempre ottenere dividendo numeratore e denominatore per b_0 , e fattorizziamo il polinomio B nella forma

$$B(x) = (x - x_1)^{k_1} \dots (x - x_r)^{k_r} ((x - \alpha_1)^2 + \beta_1^2)^{h_1} \dots ((x - \alpha_s)^2 + \beta_s^2)^{h_s}.$$

La funzione razionale si può allora decomporre nella forma

$$f(x) = \frac{A(x)}{B(x)} = \frac{a_{11}}{x - x_1} + \dots + \frac{a_{1k_1}}{(x - x_1)^{k_1}} + \dots + \frac{a_{r1}}{x - x_r} + \dots + \frac{a_{rk_r}}{(x - x_r)^{k_r}} + \dots \quad (2.11)$$

$$+ \frac{p_{11}x + q_{11}}{(x - \alpha_1)^2 + \beta_1^2} + \dots + \frac{p_{1h_1}x + q_{1h_1}}{((x - \alpha_1)^2 + \beta_1^2)^{h_1}} + \dots + \frac{p_{s1}x + q_{s1}}{(x - \alpha_s)^2 + \beta_s^2} + \dots + \frac{p_{sh_s}x + q_{sh_s}}{((x - \alpha_s)^2 + \beta_s^2)^{h_s}},$$

dove i coefficienti a_{ij} , p_{ij} , q_{ij} sono tutti da determinare. Per determinare questi coefficienti, che in tutto sono n , si moltiplicano ambo i membri di (2.11) per $B(x)$ ottenendo così un'uguaglianza tra polinomi. Dal principio di identità dei polinomi si ottiene un sistema lineare.

Passo 3. Trovati i coefficienti nello sviluppo (2.11) si può integrare ciascun pezzo mediante le tecniche che abbiamo già a disposizione.

Quanto sopra esposto può apparire complesso ma in realtà si tratta di un metodo più difficile a spiegarsi che a farsi. Qualche esempio concreto dovrebbe chiarire la situazione.

Esempio 62. Calcoliamo $\int \frac{x^2 + 2x + 3}{x(1+x^2)} dx$. Le radici del denominatore sono $x = 0$ e $x = \pm i$. Pertanto la funzione razionale si decompone in

$$\frac{x^2 + 2x + 3}{x(1+x^2)} = \frac{a}{x} + \frac{cx + d}{1+x^2}.$$

Per determinare i coefficienti giusti deve essere

$$x^2 + 2x + 3 = a(1+x^2) + (cx+d)x \Rightarrow x^2 + 2x + 3 = (a+c)x^2 + dx + a.$$

Uguagliando i coefficienti dei termini di stesso grado si ottiene

$$a + c = 1, \quad d = 2, \quad a = 3,$$

da cui $a = 3$, $d = 2$, $c = -2$. Dunque

$$\int \frac{x^2 + 2x + 3}{x(1+x^2)} dx = 3 \int \frac{1}{x} + \int \frac{2-2x}{1+x^2} dx = 3 \log|x| + 2 \arctan x - \log(1+x^2) + c.$$

Esempio 63. Calcoliamo $\int \frac{x^3}{(1+x^2)^2} dx$. Le uniche radici del denominatore sono $x = \pm i$. Allora

$$\frac{x^3}{(1+x^2)^2} = \frac{ax+b}{1+x^2} + \frac{cx+d}{(1+x^2)^2}.$$

Imponiamo la condizione $x^3 = (ax+b)(1+x^2) + (cx+d) = ax^3 + bx^2 + (a+c)x + b+d$, da cui si ricava che $a = 1$, $b = 0$, $a+c = 0$, $(b+d) = 0$. Dunque

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

Esempio 64. Calcoliamo $\int \frac{e^x - 2}{e^x + 2} dx$. Facciamo il cambiamento di variabili $e^x = t$ ottenendo

$$\int \frac{e^x - 2}{e^x + 2} dx = \int \frac{e^x - 2}{(e^x + 2)e^x} e^x dx \stackrel{e^x=t}{=} \int \frac{t-2}{(t+2)t} dt.$$

A questo punto si deve poter scrivere

$$\frac{t-2}{(t+2)t} = \frac{a}{t+2} + \frac{b}{t},$$

a patto che $t-2 = at + b(t+2) = (a+b)t + 2b$. Quindi deduciamo che $a+b = 1$ e $2b = -2$. Tornando al nostro integrale abbiamo

$$\int \frac{e^x - 2}{e^x + 2} dx \stackrel{e^x=t}{=} \int \frac{t-2}{(t+2)t} dt = 2 \int \frac{1}{t+2} dt - \int \frac{1}{t} dt = \log(t+2)^2 - \log|t| + c \stackrel{e^x=t}{=} \log(e^x+2)^2 - x + c.$$

Esempio 65.

$$\begin{aligned} \int \frac{1}{2 \sin x + \sin x \cos x} dx &= \int \frac{1}{(2 + \cos x) \sin x} dx = \\ &= \int \frac{-\sin x}{(2 + \cos x)(-\sin^2 x)} dx \stackrel{\cos x=t}{=} \int \frac{1}{(2+t)(t^2-1)} dt. \end{aligned}$$

A questo punto osserviamo che il denominatore ha le sole radici reali -2 e ± 1 . Pertanto si dovrà poter scrivere

$$\frac{1}{(t+2)(t+1)(t-1)} = \frac{a}{t+2} + \frac{b}{t+1} + \frac{c}{t-1},$$

da cui otteniamo $1 = a(t^2-1) + b(t+2)(t-1) + c(t+2)(t+1) = (a+b+c)t^2 + (b+3c)t + 2c - a - 2b$. Allora deve essere $a + b + c = 0$, $b + 3c = 0$, $2c - a - b = 1$. Pertanto si ricava che

$$a = \frac{1}{3}, \quad b = -\frac{1}{2}, \quad c = \frac{1}{6}.$$

Dunque

$$\begin{aligned} \int \frac{1}{2 \sin x + \sin x \cos x} dx &\stackrel{\cos x=t}{=} \int \frac{1}{(2+t)(t^2-1)} dt = \frac{1}{3} \log |t+2| - \frac{1}{2} \log |t+1| + \frac{1}{6} \log |t-1| + c = \\ &\stackrel{\cos x=t}{=} \log \left(\frac{\sqrt[3]{2 + \cos x} \sqrt[6]{1 - \cos x}}{\sqrt{1 + \cos x}} \right) + c. \end{aligned}$$

2.5 Integrali generalizzati

Nelle sezioni precedenti abbiamo avuto a che fare con funzioni limitate definite su di un intervallo chiuso e limitato. Ora vogliamo generalizzare la nozione di integrale in modo da contemplare situazioni in cui la funzione da integrare non sia necessariamente limitata oppure sia definita in un intervallo non limitato.

Definizione 66 (Integrale generalizzato). *Sia $f : [a, b[\rightarrow \mathbb{R}$, con b eventualmente anche $+\infty$, una funzione localmente integrabile, ovvero che sia integrabile in ogni intervallo $[a, c]$ con $c < b$. Se esiste ed è finito il limite*

$$\lim_{c \rightarrow b^-} \int_a^c f(x) dx \quad (2.12)$$

diremo che la f è integrabile in senso generalizzato, o improprio, in $[a, b[$ ed il limite (2.12) sarà indicato con la scrittura

$$\int_a^b f(x) dx.$$

Se invece il limite (2.12) esiste ed è $+\infty$ (risp. $-\infty$) diremo che l'integrale improprio diverge positivamente (risp. negativamente), e scriveremo

$$\int_a^b f(x) dx = +\infty \quad (\text{risp. } -\infty).$$

Infine, se il limite (2.12) non esiste diremo che l'integrale improprio non esiste.

Per le funzioni $f :]a, b] \rightarrow \mathbb{R}$ localmente integrabili si definisce l'integrale generalizzato in maniera del tutto analoga ponendo

$$\int_a^b f(x) dx := \lim_{c \rightarrow a^+} \int_b^c f(x) dx$$

se il limite esiste ed è finito, eccetera. Osserviamo che se f è continua su $[a, b]$ allora essa è automaticamente localmente integrabile.

Esempio 67. Applicando la definizione si riconosce che

$$\int_0^{+\infty} x dx = +\infty, \quad \int_0^{+\infty} \frac{1}{1+x^2} dx = \frac{\pi}{2}, \quad \int_0^{+\infty} \sin x dx \text{ non esiste.}$$

Definizione 68. Sia f definita su $]a, b[$ e localmente integrabile, ovvero integrabile in ogni intervallo $[\alpha, \beta]$ con $a < \alpha < \beta < b$. Scelto un punto $c \in]a, b[$, diremo che f è integrabile in senso generalizzato in $]a, b[$ se essa è integrabile in $]a, c]$ e in $[c, b[$, e in tal caso porremo

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx.$$

Se uno solo degli integrali $\int_a^c f(x) dx$ o $\int_c^b f(x) dx$ diverge positivamente (risp. negativamente), o se entrambi divergono positivamente (risp. negativamente), diremo che $\int_a^b f(x) dx$ diverge positivamente (risp. negativamente). In tutti gli altri casi diremo che $\int_a^b f(x) dx$ non esiste.

Osserviamo che la definizione precedente è ben posta, vale a dire che non dipende dalla scelta del punto $c \in]a, b[$. Infatti, se scegliessimo un ulteriore punto c' , per la proprietà di spezzamento dell'integrale avremmo che per ogni $\alpha, \beta \in]a, b[$

$$\int_\alpha^c f(x) dx + \int_c^\beta f(x) dx = \int_\alpha^\beta f(x) dx = \int_\alpha^{c'} f(x) dx + \int_{c'}^\beta f(x) dx.$$

Passando al limite otteniamo

$$\lim_{\alpha \rightarrow a^+} \int_\alpha^c f(x) dx + \lim_{\beta \rightarrow b^-} \int_c^\beta f(x) dx = \lim_{\alpha \rightarrow a^+} \int_\alpha^{c'} f(x) dx + \lim_{\beta \rightarrow b^-} \int_{c'}^\beta f(x) dx.$$

Proposizione 69. Sia f una funzione localmente integrabile e positiva in $[a, b[$. Allora l'integrale $\int_a^b f(x) dx$ esiste e può soltanto essere finito o divergere positivamente.

Dimostrazione. Consideriamo la funzione $F(x) = \int_a^x f(t) dt$, che è definita per $x < b$. Dunque dobbiamo provare che esiste, finito o infinito, il seguente limite

$$\lim_{x \rightarrow b^-} F(x).$$

Fissiamo $x < y < b$. Poiché $f \geq 0$, dalla monotonia dell'integrale deduciamo

$$\int_x^y f(t) dt \geq 0.$$

Allora

$$F(x) \leq F(x) + \int_x^y f(t) dt = \int_a^x f(t) dt + \int_x^y f(t) dt = \int_a^y f(t) dt = F(y).$$

Abbiamo dunque provato che la funzione F è una funzione positiva e crescente. L'esistenza del limite e l'ultima affermazione dell'enunciato seguono allora dalle proprietà delle funzioni monotone. \square

Naturalmente vale un enunciato corrispondente per funzioni negative o per intervalli $]a, b]$, e di conseguenza per intervalli $]a, b[$.

Esempio 70 (Esempi fondamentali). Studiamo $\int_0^1 \frac{1}{x^\alpha} dx$, che per la Proposizione 69 esiste sempre. Naturalmente se $\alpha \leq 0$ la funzione è integrabile nel senso di Riemann. Se invece $\alpha > 0$, per ogni $\varepsilon > 0$ si ha

$$\int_\varepsilon^1 \frac{1}{x^\alpha} dx = \begin{cases} -\log \varepsilon & \alpha = 1, \\ \frac{1-\varepsilon^{1-\alpha}}{1-\alpha} & \alpha \neq 1. \end{cases}$$

Passando al limite per $\varepsilon \rightarrow 0^+$ otteniamo

$$\int_0^1 \frac{1}{x^\alpha} dx < +\infty \Leftrightarrow \alpha < 1,$$

e in tal caso l'integrale vale $\frac{1}{1-\alpha}$.

Un'altro esempio fondamentale da studiare è $\int_1^{+\infty} \frac{1}{x^\alpha} dx$. Se $\alpha \leq 0$ è chiaro che l'integrale diverge positivamente. Infatti, per la monotonia dell'integrale, si ha che per ogni $c > 1$

$$\int_1^c \frac{1}{x^\alpha} dx \geq c - 1.$$

Se invece $\alpha > 1$, per ogni $c > 1$ abbiamo

$$\int_1^c \frac{1}{x^\alpha} dx = \begin{cases} \log c & \alpha = 1, \\ \frac{c^{1-\alpha}-1}{1-\alpha} & \alpha \neq 1. \end{cases}$$

Passando al limite per $c \rightarrow +\infty$ si ottiene che

$$\int_1^{+\infty} \frac{1}{x^\alpha} dx < +\infty \Leftrightarrow \alpha > 1,$$

e in tal caso l'integrale vale $\frac{1}{\alpha-1}$.

Spesso si può stabilire la convergenza di un integrale generalizzato mediante il confronto con integrali generalizzati convergenti.

Teorema 71. Siano $f, g : [a, b[\rightarrow \mathbb{R}$ due funzioni tali che

1. f è localmente integrabile,
2. $|f(x)| \leq g(x)$ per ogni $x \in [a, b[$,
3. g è integrabile in senso generalizzato in $[a, b[$.

Allora anche f risulta integrabile in senso generalizzato in $[a, b[$.

Dimostrazione. Dividiamo la funzione in parte positiva f^+ e parte negativa f^- di modo che $f = f^+ - f^-$ e $|f| = f^+ + f^-$. Essendo $f^+ \geq 0$, l'integrale $\int_a^b f^+(x) dx$ esiste. Dalla condizione 2 abbiamo che $f^+ \leq |f| \leq g$. Dalla proprietà di monotonia degli integrali e dalla condizione 3 deduciamo:

$$\int_a^b f^+(x) dx \leq \int_a^b g(x) dx < +\infty.$$

Con lo stesso ragionamento si deduce che anche f^- è integrabile in $[a, b[$. Allora è integrabile anche $f = f^+ - f^-$. \square

Osserviamo che per la validità del risultato precedente è sufficiente che le condizioni 2 e 3 siano verificate in un intorno sinistro di b . Infatti, se tali condizioni valgono in un intervallo $[\alpha, b[$, basta spezzare ogni integrale in uno su $[a, \alpha]$ e uno su $[\alpha, b[$. Naturalmente vale un risultato analogo su intervalli del tipo $]a, b]$ e $]a, b[$.

Esempio 72. La funzione

$$f(x) = \frac{1 + \sin x}{\sqrt{|x|}}$$

è integrabile in $[-1, 1]$. Infatti, separando nei due intervalli $[-1, 0[$ e $]0, 1]$ abbiamo che in ciascuno di essi risulta

$$|f(x)| \leq \frac{2}{\sqrt{|x|}},$$

e che la funzione al secondo membro della disuguaglianza è integrabile a causa del primo esempio fondamentale.

Il più delle volte il criterio del confronto si usa nella versione seguente.

Proposizione 73 (Criterio del confronto asintotico). *Siano $f, g : [a, b[\rightarrow \mathbb{R}$ due funzioni strettamente positive in un intorno di b e localmente integrabili. Supponiamo che esista*

$$L = \lim_{x \rightarrow b^-} \frac{f(x)}{g(x)}.$$

Allora si hanno:

1. se L è finito e g è integrabile allora anche f è integrabile.
2. Se $L \neq 0$ e g non è integrabile allora nemmeno f è integrabile.
3. Se L è finito e $L \neq 0$, f è integrabile se e soltanto se g è integrabile.

Dimostrazione. Supponiamo L finito e diverso da zero. Allora $|L| > 0$. Per il Teorema sulla permanenza del segno abbiamo che esistono α, β tali che in un intorno di b si verifica

$$0 < \alpha \leq \frac{|f(x)|}{|g(x)|} \leq \beta. \quad (2.13)$$

Dal Teorema 71 segue allora il punto 3. Gli altri punti seguono ugualmente usando la disuguaglianza (2.13). \square

Esempio 74. Per ogni $t > 0$ la funzione e^{-tx} è integrabile in $[a, +\infty[$. Infatti

$$\int_a^{+\infty} e^{-tx} dx = \lim_{c \rightarrow +\infty} \int_a^c e^{-tx} dx = \lim_{c \rightarrow +\infty} \left[-\frac{e^{-tx}}{t} \right]_a^c = \lim_{c \rightarrow +\infty} \left(\frac{e^{-ta} - e^{-tc}}{t} \right) = \frac{e^{-ta}}{t}.$$

Verifichiamo ora che per ogni $n \in \mathbb{N}$ si ha

$$\int_0^{+\infty} x^n e^{-x} dx < +\infty.$$

Osserviamo infatti che

$$\lim_{x \rightarrow +\infty} \frac{x^n}{e^{\frac{x}{2}}} = 0.$$

Allora

$$\lim_{x \rightarrow +\infty} \frac{x^n e^{-x}}{e^{-\frac{x}{2}}} = \lim_{x \rightarrow +\infty} \frac{x^n}{e^{\frac{x}{2}}} = 0.$$

Quindi l'integrabilità segue dal criterio del confronto asintotico e dalla prima parte dell'esempio.

Esempio 75. La funzione e^{-x^2} è integrabile in tutto \mathbb{R} . Infatti, dato che è una funzione pari possiamo limitarci a studiarla in $[0, +\infty[$. Inoltre, dal fatto che $(x - \frac{1}{2})^2 \geq 0$ deduciamo che $x^2 \geq x - \frac{1}{4}$. Pertanto

$$e^{-x^2} \leq e^{x - \frac{1}{4}} = e^{\frac{1}{4}} e^{-x}.$$

Allora l'integrabilità segue dal Teorema 71 e dall'esempio precedente.

2.6 Esercizi relativi al capitolo 1

Esercizio 76. Sia $f : [a, b] \rightarrow \mathbb{R}$ una funzione continua e positiva. Si dimostri che

$$\int_a^b f(x) dx = 0 \Rightarrow \forall x \in [a, b] : f(x) = 0.$$

Esercizio 77. Dimostrare che l'equazione

$$\int_0^x e^{-t^2} dt + x - 1 = 0$$

ammette un'unica soluzione in $[0, 1]$.

Esercizio 78. Trovare l'unica funzione derivabile $f : \mathbb{R} \rightarrow \mathbb{R}$ tale che $f(0) = 0$ e

$$f'(x) = -\frac{\sin(2x)}{1 + \sin^2 x}.$$

Esercizio 79. Calcolare l'integrale

$$\int_4^{16} \frac{1}{x \log x} dx.$$

Esercizio 80. Calcolare

$$\int_0^1 \frac{e^{2x}}{(e^x + 1)^2 (e^{2x} + 1)} dx.$$

Esercizio 81. Calcolare

$$\int \frac{\cos^2 x}{1 - 2 \sin^2 x} dx.$$

Esercizio 82. Calcolare

$$\int \frac{1}{\sin x} dx.$$

Esercizio 83. Calcolare

$$\int_1^{\sqrt{2}} \frac{x \log x}{\sqrt{x^2 + 1}} dx.$$

Esercizio 84. Si studi l'integrabilità della funzione

$$f(x) = \frac{1}{x \log^\beta x}, \quad x \in [e, +\infty[$$

al variare del parametro $\beta > 0$.

Esercizio 85. Si studi l'integrabilità della funzione

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

Esercizio 86. Calcolare

$$\int_0^{+\infty} \frac{\log x}{(x+1)^2} dx.$$

Esercizio 87. Dire se la funzione

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

è integrabile in \mathbb{R} .

Capitolo 3

Calcolo differenziale per funzioni di più variabili

In questo capitolo ci proponiamo di studiare funzioni del tipo

$$f : \mathbb{R}^N \rightarrow \mathbb{R}^M.$$

In particolare dovremo estendere a questo tipo di funzioni i concetti di continuità e derivabilità introdotti per le funzioni di una variabile reale.

3.1 Cenni di topologia di \mathbb{R}^N

Prima di tutto è necessario discutere brevemente alcune proprietà geometriche dello spazio \mathbb{R}^n . Ricordiamo la definizione di convergenza per una successione di numeri reali $(x_n) \subset \mathbb{R}$:

$$\lim_{n \rightarrow +\infty} x_n = x \Leftrightarrow \forall \varepsilon > 0 \exists \nu \geq 1 \text{ t.c. } \forall n \geq \nu, |x_n - x| < \varepsilon.$$

In altri termini, la successione converge ad un limite se da un certo indice in poi i termini della successione e il limite sono arbitrariamente ‘vicini’. Se poi $f : \mathbb{R} \rightarrow \mathbb{R}$ ed $x_0 \in \mathbb{R}$, f è continua in x_0 se

$$\forall \varepsilon > 0 \exists \delta > 0 \text{ t.c. } |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon.$$

In soldoni ciò significa che i valori della funzione sono arbitrariamente ‘vicini’ se i rispettivi punti lo sono a sufficienza. I concetti di convergenza e continuità si possono definire ogni qualvolta si sia in grado di dare un significato alla parola ‘vicino’. Questo è possibile se ad esempio si ha a disposizione una funzione distanza. Nel caso della retta reale la distanza tra due punti $x, y \in \mathbb{R}$ è data da $|x - y|$. Nel caso dello spazio \mathbb{R}^N sarà allora naturale considerare la distanza Euclidea indotta dal Teorema di Pitagora. Dunque se $x = (x_1, \dots, x_N) \in \mathbb{R}^N$ indichiamo col simbolo di norma la sua distanza dall’origine:

$$\|x\|_N := \sqrt{\sum_{i=1}^N x_i^2}. \quad (3.1)$$

Allora la distanza tra due punti $x, y \in \mathbb{R}^N$ sarà data da

$$\|x - y\|_N := \sqrt{\sum_{i=1}^N (x_i - y_i)^2}.$$

La norma soddisfa le proprietà:

1. (Omogeneità) $\|\lambda x\| = |\lambda|\|x\|$.
2. (Disuguaglianza triangolare) $\|x + y\| \leq \|x\| + \|y\|$.
3. (Positività) $\|x\| \geq 0$.
4. $\|x\| = 0 \Leftrightarrow x = 0$.

La nozione di distanza ci permette di introdurre i concetti di convergenza e continuità. Così, se x_n, x sono punti di \mathbb{R}^N diremo che

$$\lim_{n \rightarrow +\infty} x_n = x \Leftrightarrow \forall \varepsilon > 0 \exists \nu \geq 1 \text{ t.c. } \forall n \geq \nu, \|x_n - x\|_N < \varepsilon.$$

Mentre se $f: \mathbb{R}^N \rightarrow \mathbb{R}^M$, diremo che f è continua in un punto x_0 se

$$\forall \varepsilon > 0 \exists \delta > 0 \text{ t.c. } \|x - x_0\|_N < \delta \Rightarrow \|f(x) - f(x_0)\|_M < \varepsilon.$$

Spesso ometteremo l'indice nel simbolo di norma sottintendendolo. Se $x_0 \in \mathbb{R}^N$ e $r > 0$, la nozione di norma ci permette ancora di introdurre l'insieme

$$B(x_0, r) := \{x \in \mathbb{R}^N : \|x - x_0\| < r\}$$

che è detto palla aperta di centro x_0 e raggio $r > 0$.

Definizione 88 (Interno e insiemi aperti). *Sia A un sottoinsieme non vuoto di \mathbb{R}^N . Un punto $x_0 \in A$ si dice interno se esiste una palla aperta $B(x_0, r) \subset A$. L'insieme dei punti interni di A è indicato con il simbolo $\overset{\circ}{A}$. L'insieme A si dice aperto se $A = \overset{\circ}{A}$. Per definizione l'insieme vuoto si considera aperto.*

Esempio 89. *La palla aperta $B(x_0, r)$ è un insieme aperto. Infatti si consideri un qualsiasi $y \in B(x_0, r)$. Prendiamo la palla $B(y, \lambda)$ e poniamo $d = \|y - x_0\|$. Per ogni $x \in B(y, \lambda)$, usando la disuguaglianza triangolare otteniamo che*

$$\|x - x_0\| \leq \|x - y\| + \|y - x_0\| \leq \lambda + d.$$

Scegliendo $0 < \lambda < r - d$ otteniamo che $B(y, \lambda) \subset B(x_0, r)$. Pertanto, ogni punto della palla è interno e quindi la palla è aperta.

Si può verificare che $A = \overset{\circ}{A}$ è il più grande insieme aperto contenuto in A .

Esercizio 90. *Si verifichi che*

1. \emptyset, \mathbb{R}^N sono aperti.
2. L'unione qualsiasi di insiemi aperti è ancora un insieme aperto.
3. L'intersezione finita di insiemi aperti è ancora un insieme aperto.

Definizione 91 (Insiemi chiusi e chiusura). *Un sottoinsieme C di \mathbb{R}^N si dice chiuso se il suo complementare è aperto. Si dice chiusura di A , e si indica con \bar{A} , il più piccolo insieme chiuso contenente A .*

E' immediato verificare che A è chiuso se e soltanto se $A = \bar{A}$. Inoltre segue immediatamente che

$$A \subset B \Rightarrow \bar{A} \subset \bar{B}.$$

Esercizio 92. *Si verifichi che*

1. \emptyset, \mathbb{R}^N sono chiusi.
2. L'intersezione qualsiasi di insiemi chiusi è ancora un insieme chiuso.
3. L'unione finita di insiemi chiusi è ancora un insieme chiuso.

Le 3 proprietà dell'esercizio 90, o 92, definiscono quella che si chiama una topologia su \mathbb{R}^N . I concetti di convergenza e continuità sono essenzialmente dei concetti topologici in quanto si possono caratterizzare in termini di insiemi aperti e chiusi.

La chiusura di un insieme si può caratterizzare in termini di convergenza

Proposizione 93. *Le seguenti proprietà sono equivalenti:*

1. $x \in \overline{A}$,
2. $\forall r > 0, A \cap B(x, r) \neq \emptyset$,
3. $\exists (x_n) \subset A$ t.c. $x_n \rightarrow x$, per $n \rightarrow +\infty$.

Dimostrazione. Sia $x \in \overline{A}$. Se per assurdo $A \cap B(x, r) = \emptyset$, per qualche $r > 0$, allora avremmo che $A \subset B(x, r)^c$ e quindi la contraddizione $x \in \overline{A} \subset B(x, r)^c$. Sia ora vera la 2). Allora, per ogni $n \geq 1$ possiamo scegliere

$$x_n \in A \cap B(x, 1/n).$$

E' immediato allora verificare che $x_n \rightarrow x$, per $n \rightarrow +\infty$. Infine, sia vera la 3). Se per assurdo fosse $x \in \overline{A}^c$ allora, essendo un insieme aperto, sarebbe $B(x, r) \subset \overline{A}^c$ per qualche $r > 0$. Allora si avrebbe $B(x, r) \cap A = \emptyset$. Poichè $x_n \rightarrow x$, si può trovare un indice $n \geq 1$ tale che $\|x_n - x\| \leq r$ da cui si otterrebbe la contraddizione $x_n \in A \cap B(x, r)$. \square

Definizione 94 (Frontiera). *Si dice frontiera, o bordo, di A l'insieme*

$$F(A) = \overline{A} \setminus \overset{\circ}{A}.$$

Esercizio 95. *Si verifichi che la frontiera è un insieme chiuso. Inoltre si verifichi che*

$$\overline{B(x_0, r)} = \{x \in \mathbb{R}^N : \|x - x_0\| \leq r\}, \quad F(B(x_0, r)) = \{x \in \mathbb{R}^N : \|x - x_0\| = r\},$$

e quindi che la palla chiusa $\overline{B}(x_0, r) = \{x \in \mathbb{R}^N : \|x - x_0\| \leq r\}$ è un insieme chiuso.

Definizione 96 (Limitatezza e compattezza). *Un insieme A si dice limitato se esiste una palla $B(0, r)$ tale che $A \subset B(0, r)$. Dunque se esiste un numero $M > 0$ tale che*

$$\forall x \in A, \|x\| \leq M.$$

Diremo poi che A è compatto se è chiuso e limitato.

Analogamente a quanto visto per le funzioni di una variabile valgono i seguenti risultati.

Teorema 97 (di Weierstrass). *Sia $f : A \subset \mathbb{R}^N \rightarrow \mathbb{R}$ una funzione continua. Se A è compatto allora la f è dotata di massimo e di minimo.*

Teorema 98 (di Heine-Cantor). *Sia $f : A \subset \mathbb{R}^N \rightarrow \mathbb{R}^M$ una funzione continua. Se A è compatto allora la f è uniformemente continua.*

3.2 Richiami di algebra lineare

Una prima classe importante di funzioni di più variabili è costituita dalle applicazioni lineari. Cominciamo con l'introdurre alcuni concetti di algebra lineare che ci saranno utili in seguito. Il lettore interessato a maggiori dettagli potrà consultare ad esempio [2] o anche [3]. Intanto ricordiamo che lo spazio \mathbb{R}^N è dotato di una struttura di spazio vettoriale in quanto è possibile definire la somma e il prodotto per uno scalare mediante le leggi

$$\forall x, y \in \mathbb{R}^N, \quad x + y := (x_1 + y_1, \dots, x_N + y_N), \quad \forall x \in \mathbb{R}^N, \forall \lambda \in \mathbb{R}, \quad \lambda x := (\lambda x_1, \dots, \lambda x_N),$$

che godono delle usuali proprietà algebriche. Spesso chiameremo dunque i punti di \mathbb{R}^N vettori.

Definizione 99. *Dati v_1, \dots, v_r r -vettori di uno spazio vettoriale V , si dice che un vettore $v \in V$ è combinazione lineare dei vettori v_1, \dots, v_r se esistono r -scalari $\lambda_1, \dots, \lambda_r \in \mathbb{R}$ tali che*

$$v = \sum_{i=1}^r \lambda_i v_i.$$

Si dice spazio generato dai vettori v_1, \dots, v_r , e lo indicheremo con $L(v_1, \dots, v_r)$, l'insieme di tutte le combinazioni lineari dei vettori v_1, \dots, v_r .

Definizione 100 (Lineare indipendenza). *I vettori $v_1, \dots, v_r \in V$ si dicono linearmente indipendenti se comunque si prendano $\lambda_1, \dots, \lambda_r \in \mathbb{R}$ si ha*

$$\sum_{i=1}^r \lambda_i v_i = 0 \Rightarrow \lambda_1 = \dots = \lambda_r = 0.$$

In caso contrario, se cioè la combinazione lineare $\sum_{i=1}^r \lambda_i v_i$ si può annullare anche se qualcuno dei $\lambda_1, \dots, \lambda_r$ non è nullo, diremo che i vettori $v_1, \dots, v_r \in V$ sono linearmente dipendenti.

Definizione 101 (Dimensione). *Uno spazio vettoriale V si dice di dimensione finita se esiste un sistema finito di vettori (v_1, \dots, v_r) tali che $L(v_1, \dots, v_r) = V$. Un sistema di vettori linearmente indipendenti tali che $L(v_1, \dots, v_r) = V$ si dice base. Ogni spazio di dimensione finita ammette almeno una base e tutte le basi hanno lo stesso numero di elementi. Tale numero è detto dimensione dello spazio.*

La dimensione finita di uno spazio vettoriale si può anche caratterizzare come il massimo numero possibile di vettori linearmente indipendenti.

Esempio 102. *Consideriamo in \mathbb{R}^2 i vettori $e_1 = (1, 0), e_2 = (0, 1)$. Tali vettori sono linearmente indipendenti. Infatti se prendiamo due qualsiasi numeri $\lambda, \mu \in \mathbb{R}$ abbiamo*

$$\lambda e_1 + \mu e_2 = 0 \Rightarrow (\lambda, 0) + (0, \mu) = 0 \Rightarrow (\lambda, \mu) = 0 \Rightarrow \lambda = \mu = 0.$$

Inoltre se $x = (x_1, x_2)$ è un qualsiasi vettore di \mathbb{R}^2 abbiamo

$$x = (x_1, x_2) = x_1(1, 0) + x_2(0, 1) = x_1 e_1 + x_2 e_2.$$

Dunque $L(e_1, e_2) = \mathbb{R}^2$. Dunque i vettori (e_1, e_2) formano una base di \mathbb{R}^2 detta base canonica. Allo stesso modo si può verificare che i vettori

$$e_1 = (1, 0, \dots, 0), \quad e_2 = (0, 1, \dots, 0), \quad \dots, \quad e_N = (0, 0, \dots, 1),$$

formano una base di \mathbb{R}^N detta base canonica.

Se v_1, \dots, v_r è una base di V , essendo $L(v_1, \dots, v_r) = V$, ogni vettore $v \in V$ si può esprimere come combinazione lineare degli elementi della base. Ovvero

$$v = \sum_{i=1}^r \lambda_i v_i. \quad (3.2)$$

L'espressione in (3.2) è in realtà l'unica possibile e i coefficienti λ_i vengono detti coordinate del vettore v rispetto alla base (v_1, \dots, v_r) . Per ogni $x = (x_1, \dots, x_N) \in \mathbb{R}^N$ le coordinate x_i sono riferite naturalmente alla base canonica di \mathbb{R}^N .

3.2.1 Matrici e applicazioni lineari

Siano V e V' due spazi vettoriali di dimensione rispettivamente N e M . Siano (e_1, \dots, e_N) e (e'_1, \dots, e'_M) rispettivamente due basi di V e V' .

Definizione 103 (Applicazioni lineari). *Una funzione $f : V \rightarrow V'$ si dice lineare se per ogni $u, v \in V$ e per ogni $\lambda, \mu \in \mathbb{R}$ risulta*

$$f(\lambda u + \mu v) = \lambda f(u) + \mu f(v).$$

Si indicherà con $\mathcal{L}(V, V')$ l'insieme delle applicazioni lineari di V in V' .

Osserviamo che se f è lineare allora $f(0) = 0$. Inoltre, si può verificare facilmente che la condizione di linearità vale per ogni combinazione lineare finita, ovvero

$$f\left(\sum_{i=1}^r \lambda_i u_i\right) = \sum_{i=1}^r \lambda_i f(u_i).$$

Pertanto, ogni $f \in \mathcal{L}(V, V')$ sarà univocamente determinata dal valore che la f assume sui vettori della base. Infatti se $u \in V$, indicate con u_i le coordinate di u avremo che

$$f(u) = f\left(\sum_{i=1}^N u_i e_i\right) = \sum_{i=1}^N u_i f(e_i),$$

e quindi è sufficiente conoscere i valori $f(e_i)$ per individuare la funzione lineare f per intero. Ora, $f(e_i) \in V'$ e quindi avrà diritto alle sue coordinate, diciamo a_{ij} , $j = 1, \dots, M$ per cui

$$f(e_i) = \sum_{j=1}^M a_{ij} e'_j. \quad (3.3)$$

Restano così individuati $N \times M$ numeri reali $\{a_{ij} \in \mathbb{R} : i = 1, \dots, N; j = 1, \dots, M\}$, o come si suol dire una matrice

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1M} \\ a_{21} & a_{22} & \dots & a_{2M} \\ \dots & \dots & \dots & \dots \\ a_{N1} & a_{N2} & \dots & a_{NM} \end{pmatrix}$$

ad N righe e M colonne. Nel seguito scriveremo brevemente

$$A = (a_{ij})_{\substack{i=1, \dots, N \\ j=1, \dots, M}}$$

dove i, j denotano rispettivamente l'indice di riga e quello di colonna. Spesso scriveremo semplicemente $A = (a_{ij})$ ogni volta che gli indici sono chiari dal contesto. Indicheremo poi con il simbolo $\mathcal{M}(N, M)$ l'insieme delle matrici ad N righe e M colonne e con $\mathcal{M}(N)$ l'insieme delle

matrici quadrate di ordine N , ovvero quelle con N righe ed N colonne. Osserviamo che $\mathcal{M}(N, M)$ è dotato di una struttura di spazio vettoriale. Se infatti $A = (a_{ij}), B = (b_{ij}) \in \mathcal{M}(N, M)$ e $\lambda \in \mathbb{R}$, si definiscono le matrici somma e prodotto per uno scalare mediante:

$$A + B = (c_{ij}), \quad c_{ij} = a_{ij} + b_{ij}; \quad \lambda A = (d_{ij}), \quad d_{ij} = \lambda a_{ij}.$$

Se le dimensioni della matrice sono opportune si può definire l'importante nozione di prodotto tra matrici.

Definizione 104 (Prodotto righe per colonne). *Siano $A = (a_{ij}) \in \mathcal{M}(N, M)$ e $B = (b_{jk}) \in \mathcal{M}(M, P)$. Allora si definisce la matrice $AB \in \mathcal{M}(N, P)$ mediante:*

$$AB = (c_{ik})_{\substack{i=1, \dots, N \\ k=1, \dots, P}} \quad c_{ik} = \sum_{j=1}^M a_{ij} b_{jk}.$$

Osserviamo che l'espressione $f(e_i) = \sum_{j=1}^M a_{ij} e'_j$ in (3.3) si può allora esprimere in forma compatta mediante il prodotto righe per colonne $f(e_i) = A e'_j$ con $A = (a_{ij})$. Se dunque consideriamo ad esempio il caso $V = \mathbb{R}^N$ e $V' = \mathbb{R}^M$, ad ogni matrice $A \in \mathcal{M}(M, N)$ corrisponde un'applicazione lineare $f: \mathbb{R}^N \rightarrow \mathbb{R}^M$ e viceversa ad ogni applicazione lineare corrisponde una certa matrice.

Alle matrici quadrate si può associare un numero reale $\det(A)$ detto determinante che gioca un ruolo importante in molte questioni. Per le matrici di ordine due esso è la differenza tra il prodotto della diagonale principale e di quella secondaria ovvero

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

Nel caso di matrici generiche $A = (a_{ij}) \in \mathcal{M}(N)$ esso può essere calcolato ad esempio con la regola di Laplace che consiste dapprima nello scegliere una riga, o una colonna (in genere si sceglie quella che contiene il maggior numero di zeri), diciamo la r -esima. Allora

$$\det(A) = \sum_{i=1}^N a_{ri} M_{ri} = \sum_{i=1}^N a_{ir} M_{ir},$$

dove $M_{ri} = (-1)^{r+i} \det(A_{ri})$ con A_{ri} la matrice che si ottiene da A sopprimendo la r -esima riga e la i -esima colonna.

La matrice $I \in \mathcal{M}(N)$ definita da

$$I = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

ovvero quella che ha termini non nulli e tutti pari ad 1 solo e soltanto lungo la diagonale principale è detta matrice identità. In effetti si verifica:

$$\forall A \in \mathcal{M}(N), \quad AI = A = IA.$$

Proposizione 105 (Proprietà del determinante). *Siano $v_1, \dots, v_N \in \mathbb{R}^N$ e $A = (v_1, \dots, v_N) \in \mathcal{M}(N)$ la matrice con le colonne costituite dai vettori $v_i, i = 1, \dots, N$. Allora, per ogni $u_i, v_i \in \mathbb{R}^N$ e per ogni $\lambda, \mu \in \mathbb{R}$ si ha*

1. (linearità)

$$\det(v_1, \dots, \lambda u_i + \mu v_i, \dots, v_N) = \lambda \det(v_1, \dots, u_i, \dots, v_N) + \mu \det(v_1, \dots, v_i, \dots, v_N),$$

2. (alternanza) $\det(v_1, \dots, v_i, \dots, v_j, \dots, v_N) = -\det(v_1, \dots, v_j, \dots, v_i, \dots, v_N)$,
3. $\det(v_1, \dots, v_i, \dots, \lambda v_i, \dots, v_N) = 0$,
4. (Cauchy-Binet) per ogni $A, B \in \mathcal{M}(N)$ si ha $\det(AB) = \det(A)\det(B)$,
5. $\det(I) = 1$.

Una matrice $A \in \mathcal{M}(N)$ si dice invertibile se esiste una matrice A^{-1} tale che

$$AA^{-1} = I = A^{-1}A.$$

Qualora esista, la matrice inversa è unica.

Teorema 106. *Una matrice quadrata è invertibile se e soltanto se il suo determinante è diverso da zero.*

Dimostrazione. Supponiamo che A sia invertibile. Allora

$$AA^{-1} = I \Rightarrow 1 = \det(AA^{-1}) = \det(A)\det(A^{-1}) \Rightarrow \det(A) \neq 0.$$

Viceversa, se $\det(A) \neq 0$, definiamo la matrice $A^{-1} = (b_{ij})$ dove

$$b_{ij} = \frac{M_{ji}}{\det(A)}. \quad (3.4)$$

Dunque abbiamo che

$$(AA^{-1})_{ij} = \sum_{h=1}^N a_{ih}b_{hj} = \frac{1}{\det(A)} \sum_{h=1}^N a_{ih}M_{jh}.$$

Ora osserviamo che la quantità $\sum_{h=1}^N a_{ih}M_{jh}$, vale $\det(A)$ se $i = j$ (regola di Laplace), mentre vale 0 se $i \neq j$, poichè in tal caso corrisponde al determinante della matrice che si ottiene da A sostituendo la j -esima riga con la i -esima. \square

La nozione di determinante è utile anche per determinare la lineare dipendenza di vettori.

Teorema 107. *Dati $v_1, \dots, v_N \in \mathbb{R}^N$ essi sono linearmente indipendenti se e soltanto se la matrice che ha per colonne i -esime le coordinate di v_i ha determinante diverso da zero.*

Come è noto, lo spazio \mathbb{R}^N è anche dotato di un prodotto scalare tra vettori definito da

$$\langle x, y \rangle := \sum_{i=1}^N x_i y_i.$$

Si dice allora che la norma definita in (3.1) è indotta dal prodotto scalare in quanto

$$\|x\| = \sqrt{\langle x, x \rangle}.$$

Esercizio 108. *Il prodotto scalare verifica le proprietà*

1. (Simmetria) $\langle x, y \rangle = \langle y, x \rangle$.
2. (Linearità) $\langle \lambda x + \mu y, z \rangle = \lambda \langle x, z \rangle + \mu \langle y, z \rangle$.
3. (Positività) $\langle x, x \rangle \geq 0$.
4. $\langle x, x \rangle = 0 \Leftrightarrow x = 0$.

Importante è anche la seguente disuguaglianza di Cauchy-Schwarz

$$|\langle x, y \rangle| \leq \|x\| \|y\|.$$

Definizione 109 (Sistemi ortonormali). *Due vettori $x, y \in \mathbb{R}^N$ si dicono ortogonali se $\langle x, y \rangle = 0$. Un sistema di vettori v_1, \dots, v_r si dice ortonormale se i vettori sono a due a due ortogonali tra loro e $\|v_i\| = 1$.*

Evidentemente, la base canonica (e_1, \dots, e_N) è un sistema ortonormale di \mathbb{R}^N .

Proposizione 110. *I vettori di un sistema ortonormale sono linearmente indipendenti.*

Dimostrazione. Siano v_1, \dots, v_r i vettori del sistema ortonormale. Consideriamo allora una qualsiasi combinazione lineare dei vettori che sia nulla:

$$\sum_{i=1}^r \lambda_i v_i = 0.$$

Fissiamo un indice $j = 1, \dots, r$ e moltiplichiamo scalarmente per il vettore v_j . Dalle proprietà del prodotto scalare otteniamo

$$0 = \langle v_j, \sum_{i=1}^r \lambda_i v_i \rangle = \sum_{i=1}^r \langle v_j, \lambda_i v_i \rangle = \sum_{i=1}^r \lambda_i \langle v_j, v_i \rangle = \lambda_j.$$

□

Definizione 111 (Autovalori e autovettori). *Sia $A \in \mathcal{M}(N)$. Si dice che $\lambda \in \mathbb{R}$ è un autovalore della matrice A se esiste un vettore $v \in \mathbb{R}^N$ diverso da zero tale che*

$$Av = \lambda v. \quad (3.5)$$

Il vettore v che soddisfa la (3.5) viene detto autovettore corrispondente all'autovalore λ .

Osserviamo che se v è un autovettore allora ogni suo multiplo αv è un autovettore per ogni $\alpha \neq 0$. Inoltre si verifica immediatamente che se v, w sono due autovettori corrispondenti all'autovalore λ , allora anche la loro somma $v + w$ è un autovalore. Dunque l'insieme degli autovalori costituisce uno spazio vettoriale, o meglio un sottospazio vettoriale di \mathbb{R}^N . La relazione (3.5) può essere scritta anche nella forma

$$(A - \lambda I)v = 0.$$

Pertanto, dire che v è un autovettore di A equivale a chiedere che la matrice $A - \lambda I$ non sia invertibile (altrimenti avremmo $v = 0$). Tenuto conto del Teorema 3.4 abbiamo che $\lambda \in \mathbb{R}$ è un autovalore se e soltanto se $\det(A - \lambda I) = 0$, ovvero se λ è una soluzione reale dell'equazione caratteristica

$$\det(A - \lambda I) = 0.$$

Poichè $A \in \mathcal{M}(N)$, l'equazione caratteristica è una equazione algebrica di grado N nell'incognita λ . Dunque la matrice A ammette al più N autovalori. L'esistenza di autovalori è garantita se la matrice A è simmetrica.

Definizione 112 (Matrici simmetriche). *Data una matrice $A = (a_{ij})$ si dice matrice trasposta di A la matrice $A^t = (a_{ji})$ che si ottiene da A scambiando le righe con le colonne. Una matrice A si dice simmetrica se $A = A^t$, ovvero se $a_{ij} = a_{ji}$.*

Dalla regola di Laplace segue subito che $\det(A) = \det(A^t)$.

Teorema 113. *Sia $A \in \mathcal{M}(N)$ una matrice simmetrica. Allora tutte le soluzioni dell'equazione caratteristica sono reali.*

Dunque una matrice simmetrica ammette almeno un autovalore. Abbiamo osservato che gli autovettori formano un sottospazio vettoriale la cui dimensione si può dimostrare essere uguale alla molteplicità algebrica dell'autovalore.

Una matrice $A \in \mathcal{M}(N)$ è legata alla sua trasposta dalla fondamentale relazione

$$\forall x, y \in \mathbb{R}^N, \langle Ax, y \rangle = \langle x, A^t y \rangle. \quad (3.6)$$

Infatti, se $A = (a_{ij})$ abbiamo che

$$\begin{aligned} \langle Ax, y \rangle &= \sum_{k=1}^N \left(\sum_{i=1}^N a_{ki} x_i \right) y_k = \sum_{k=1}^N \left(\sum_{i=1}^N a_{ki} x_i y_k \right) = \sum_{i=1}^N \left(\sum_{k=1}^N a_{ki} x_i y_k \right) = \\ &= \sum_{i=1}^N \left(\sum_{k=1}^N a_{ki} y_k \right) x_i = \sum_{i=1}^N (A^t y)_i x_i = \langle A^t y, x \rangle. \end{aligned}$$

La simmetria della matrice è legata anche alla lineare indipendenza degli autovalori.

Teorema 114. *Sia $A \in \mathcal{M}(N)$ una matrice simmetrica. Allora ad autovalori distinti corrispondono autovettori linearmente indipendenti.*

Dimostrazione. Siano $\lambda, \mu \in \mathbb{R}$ due autovalori distinti. Consideriamo allora due autovettori v, w corrispondenti a λ e μ . Dunque

$$Av = \lambda v, \quad Aw = \mu w.$$

Allora

$$\lambda \langle v, w \rangle = \langle Av, w \rangle = \langle v, Aw \rangle = \mu \langle v, w \rangle,$$

da cui otteniamo che

$$(\lambda - \mu) \langle v, w \rangle = 0.$$

Poichè $\lambda \neq \mu$ deduciamo che i vettori v, w sono ortogonali. Allora per la Proposizione 110 essi sono linearmente indipendenti. \square

3.3 Sistemi lineari

Se $A \in \mathcal{M}(M, N)$, e $b \in \mathbb{R}^M$, si dice sistema lineare l'equazione

$$Ax = b. \quad (3.7)$$

Ogni $x \in \mathbb{R}^N$ che soddisfa (3.7) si dice soluzione del sistema lineare. Se $b = 0$ il sistema si dice omogeneo. Un sistema si dice compatibile o incompatibile a seconda che ammetta o meno soluzioni. La matrice A è detta matrice dei coefficienti.

Teorema 115 (di Cramer). *Sia $A \in \mathcal{M}(N)$. Allora il sistema (3.7) ammette una unica soluzione quale che sia $b \in \mathbb{R}^N$ se e soltanto se $\det(A) \neq 0$.*

Naturalmente, tale soluzione è data da $x = A^{-1}b$. Inoltre, se $A = (v_1, \dots, v_N) \in \mathcal{M}(N)$, indicata con A_i la matrice ottenuta da A sostituendo il vettore v_i con il vettore b , la soluzione può essere calcolata tramite la formula di Cramer

$$x_i = \frac{\det(A_i)}{\det(A)}, \quad \forall i = 1, \dots, N.$$

Nelle ipotesi del Teorema di Cramer il sistema omogeneo ammette la sola soluzione nulla.

Definizione 116. Sia $A \in \mathcal{M}(M, N)$. Una sottomatrice $p \times q$ di A è la matrice costituita da elementi comuni a p righe e q colonne. Si dice rango di A , o caratteristica, il numero massimo degli ordini delle sue sottomatrici quadrate invertibili, ovvero con determinante diverso da zero. Indicheremo tale numero con il simbolo $r(A)$.

Dunque $r(A) \leq \min(M, N)$.

Teorema 117 (Rouchè-Capelli). Sia $A \in \mathcal{M}(M, N)$, $b \in \mathbb{R}^M$. Considerata la matrice completa \hat{A} ottenuta da A aggiungendo il vettore colonna b , il sistema (3.7) ammette soluzione se e soltanto se $r(\hat{A}) = r(A)$.

3.3.1 Il metodo di eliminazione di Gauss

I risultati precedentemente esposti sui sistemi lineari sono poco adatti alle esigenze numeriche. Ad esempio può non essere molto agevole stabilire al calcolatore se effettivamente il determinante di una matrice è nullo. In effetti dalla relazione $\det(\lambda A) = \lambda^N \det(A)$, se ad esempio $\det(A) = 1$ e $N = 30$ abbiamo che $\det(10^{-1}A) = 10^{-30}$.

Dunque, sarà utile introdurre una procedura algoritmica che ci permetta di stabilire se un sistema è compatibile o no e allo stesso tempo consenta di calcolare le soluzioni. Uno di questi procedimenti è il metodo di eliminazione di Gauss.

Cominciamo con il considerare un sistema lineare a *gradini*, ovvero un sistema lineare di m equazioni nelle incognite x_1, \dots, x_n con $m \leq n$ della forma

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{mm}x_m + \dots + a_{mn}x_n &= b_m \end{aligned} \quad (3.8)$$

con $a_{11}a_{22} \dots a_{mm} \neq 0$. La matrice dei coefficienti è del tipo

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & \dots & \dots & a_{1n} \\ 0 & a_{22} & \dots & \dots & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & a_{mm} & \dots & a_{mn} \end{pmatrix}.$$

Nel caso $m = n$, l'ultima equazione di (3.8) è soddisfatta dal solo valore $x_n = b_n/a_{nn}$. Sostituendo tale valore nella penultima equazione si trova un unico valore per x_{n-1} . I valori x_n, x_{n-1} così ottenuti, sostituiti nella terz'ultima equazione danno luogo ad un unico valore x_{n-2} . Procedendo in questo modo si ottiene un'unica soluzione di (3.8). Quindi *un sistema a gradini di n equazioni in n incognite è sempre compatibile ed ammette un'unica soluzione.*

Se $m < n$ il sistema (3.8) può essere scritto nella forma equivalente:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1m}x_m &= b_1 - (a_{1m+1}x_{m+1} + \dots + a_{1n}x_n) \\ a_{22}x_2 + \dots + a_{2m}x_m &= b_2 - (a_{2m+1}x_{m+1} + \dots + a_{2n}x_n) \\ &\vdots \\ a_{mm}x_m &= b_m - (a_{mm+1}x_{m+1} + \dots + a_{mn}x_n). \end{aligned} \quad (3.9)$$

Dando valori arbitrari $t_{m+1}, \dots, t_n \in \mathbb{R}$ alle incognite x_{m+1}, \dots, x_n si ottiene un sistema a gradini di

m equazioni in m incognite

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m &= b_1 - (a_{1m+1}t_{m+1} + \cdots + a_{1n}t_n) \\ a_{22}x_2 + \cdots + a_{2m}x_m &= b_2 - (a_{2m+1}t_{m+1} + \cdots + a_{2n}t_n) \\ &\vdots \\ a_{mm}x_m &= b_m - (a_{mm+1}t_{m+1} + \cdots + a_{mn}t_n) \end{aligned} \quad (3.10)$$

che ammette un' unica soluzione. Pertanto il sistema (3.8) di partenza ammette infinite soluzioni al variare dei parametri $t_{m+1}, \dots, t_n \in \mathbb{R}$. In particolare esprimeremo questo fatto dicendo che il sistema (3.8) ammette ∞^{n-m} soluzioni. Ora, dal modo in cui si calcolano tali soluzioni si deduce che ogni soluzione di (3.8) si esprime mediante una n -pla

$$(S_1(t_{m+1}, \dots, t_n), S_2(t_{m+1}, \dots, t_n), \dots, S_n(t_{m+1}, \dots, t_n)) \quad (3.11)$$

in cui gli S_i sono polinomi di primo grado nei parametri t_{m+1}, \dots, t_n . La (3.11) si dice soluzione generale del sistema (3.8). In particolare, vediamo che un sistema a gradini è sempre compatibile.

Due sistemi lineari nelle incognite x_1, \dots, x_n con $m \leq n$ si dicono *equivalenti* se possiedono le stesse soluzioni. Il metodo di eliminazione di Gauss consiste nel sostituire il sistema assegnato con sistemi ad esso equivalenti mediante passaggi successivi fino a raggiungere la forma a gradini. Possiamo identificare le seguenti operazioni elementari sulle equazioni di un sistema che conducono a sistemi equivalenti:

1. Scambiare tra loro due equazioni del sistema.
2. Moltiplicare un' equazione per uno scalare non nullo.
3. Sostituire una equazione con quella ottenuta sommando ad essa un multiplo di un'altra equazione.

A queste operazioni corrispondono naturalmente altrettante operazioni elementari sulla matrice completa del sistema:

1. Scambiare tra loro due righe della matrice.
2. Moltiplicare una riga della matrice per uno scalare non nullo.
3. Sostituire una riga della matrice con quella ottenuta sommando ad essa un multiplo di un'altra riga.

Sia dunque assegnato il sistema

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m. \end{aligned} \quad (3.12)$$

Se una delle sue equazioni è del tipo $0 = b_i$ allora essa sarà soddisfatta se $b_i = 0$ e in tal caso potremo cancellarla dal sistema, altrimenti il sistema è incompatibile. Dunque possiamo supporre che tutti i primi membri delle equazioni del sistema siano non nulli. Inoltre, possiamo certamente assumere che $a_{i1} \neq 0$ per qualche $i = 1, \dots, m$. Ciò può essere sempre fatto scambiando eventualmente due incognite tra loro. Con l'operazione 1) possiamo ottenere $a_{11} \neq 0$. Con l'operazione 2) possiamo poi

assumere che $a_{11} = 1$. Sommando alle successive equazioni la prima moltiplicata rispettivamente per $-a_{21}, -a_{31}, \dots, a_{m1}$ (operazione 3)) si ottiene un nuovo sistema

$$\begin{aligned} x_1 + a'_{12}x_2 + \dots + a'_{1n}x_n &= b'_1 \\ a'_{22}x_2 + \dots + a'_{2n}x_n &= b'_2 \\ &\vdots \\ a'_{m2}x_2 + \dots + a'_{mn}x_n &= b'_m \end{aligned} \quad (3.13)$$

Se qualcuna delle equazioni del sistema (3.13) è del tipo $0 = 0$ la omettiamo. Se invece compare un'equazione della forma $0 = b'_i$ con $b'_i \neq 0$ allora il sistema è incompatibile e il procedimento si arresta. Dunque possiamo supporre che tutti i primi membri del sistema (3.13) siano non nulli. A questo punto procediamo sul sistema senza più considerare la prima equazione e operando sulle rimanenti equazioni come nel caso precedente. Tale procedura può così essere iterata fino a che non si ottiene l'incompatibilità del sistema o la sua riduzione in forma a gradino, da cui si ottiene la soluzione generale del sistema. Nella pratica è preferibile operare sulla matrice completa del sistema piuttosto che sulle equazioni.

Esercizio 118.

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 1 \\ 2x_1 + x_2 + 4x_3 &= 2 \\ 3x_1 - 3x_2 + x_3 &= 1. \end{aligned}$$

Eseguiamo le operazioni elementari sulle righe della matrice completa:

$$\begin{aligned} \left(\begin{array}{cccc} 1 & 2 & 3 & 1 \\ 2 & 1 & 4 & 2 \\ 3 & -3 & 1 & 1 \end{array} \right) &\rightarrow \left(\begin{array}{cccc} 1 & 2 & 3 & 1 \\ 0 & -3 & -2 & 0 \\ 0 & -9 & -8 & -2 \end{array} \right) \rightarrow \\ &\rightarrow \left(\begin{array}{cccc} 1 & 2 & 3 & 1 \\ 0 & 1 & \frac{2}{3} & 0 \\ 0 & 0 & -2 & -2 \end{array} \right) \rightarrow \left(\begin{array}{cccc} 1 & 2 & 3 & 1 \\ 0 & -3 & -2 & 0 \\ 0 & 0 & 1 & 1 \end{array} \right). \end{aligned}$$

Il sistema a gradini è il seguente

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 1 \\ x_2 + \frac{2}{3}x_3 &= 0 \\ x_3 &= 1, \end{aligned}$$

che ammette l'unica soluzione $(-\frac{2}{3}, -\frac{2}{3}, 1)$.

Esercizio 119.

$$\begin{aligned} x_3 + 2x_4 &= 3 \\ 2x_1 + 4x_2 - 2x_3 &= 4 \\ 2x_1 + 4x_2 - x_3 + 2x_4 &= 7. \end{aligned}$$

Operando sulle righe della matrice completa abbiamo

$$\left(\begin{array}{ccccc} 0 & 0 & 1 & 2 & 3 \\ 2 & 4 & -2 & 0 & 4 \\ 2 & 4 & -1 & 2 & 7 \end{array} \right) \rightarrow \left(\begin{array}{ccccc} 1 & 2 & -1 & 0 & 2 \\ 2 & 4 & -1 & 2 & 7 \\ 0 & 0 & 1 & 2 & 3 \end{array} \right) \rightarrow$$

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

Il sistema corrispondente è

$$\begin{aligned} x_1 + 2x_2 - x_3 &= 2 \\ x_3 + 2x_4 &= 3. \end{aligned}$$

Scambiando x_2 con x_3 otteniamo il sistema a gradini

$$\begin{aligned} x_1 - x_3 + 2x_2 &= 2 \\ x_3 + 2x_4 &= 3. \end{aligned}$$

Pertanto la soluzione generale del sistema è data da

$$(x_1, x_2, x_3, x_4) = (5 - 2t - 2s, t, 3 - 2s, s), \quad t, s \in \mathbb{R}.$$

Esercizio 120.

$$\begin{aligned} x_2 - x_3 &= -1 \\ x_1 + x_3 &= 1 \\ 2x_1 + x_2 + x_3 &= 2. \end{aligned}$$

Operando sulle righe della matrice completa abbiamo

$$\begin{aligned} &\begin{pmatrix} 0 & 1 & -1 & -1 \\ 1 & 0 & 1 & 1 \\ 2 & 1 & 1 & 2 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & -1 & -1 \\ 2 & 1 & 1 & 2 \end{pmatrix} \rightarrow \\ &\rightarrow \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & -1 & -1 \\ 0 & 1 & -1 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

La terza riga corrisponde all'equazione incompatibile $0 = 1$. Il sistema pertanto è incompatibile.

3.4 Derivate parziali e differenziabilità

In questa sezione cominciamo a studiare la teoria della derivazione per funzioni di più variabili. Per quanto tale teoria presenti larghe analogie con quella delle funzioni reali di una variabile, essa presenta anche nuove caratteristiche che dipendono essenzialmente dal fatto che in dimensione maggiore di uno ci sono più direzioni lungo le quali ci si può avvicinare ad un punto.

3.4.1 Derivate parziali

Sia $A \subset \mathbb{R}^2$ un insieme aperto ed $f : A \rightarrow \mathbb{R}$. Fissato un punto $(x_0, y_0) \in A$ esisterà una piccola palla e in particolare un piccolo quadrato contenuto in A . Allora ha senso considerare la funzione parziale $f(\cdot, y_0)$ definita da

$$x \in]x_0 - \delta, x_0 + \delta[\mapsto f(x, y_0) \in \mathbb{R}$$

per $\delta > 0$ abbastanza piccolo. Diremo che la f ammette derivata parziale rispetto alla prima variabile nel punto (x_0, y_0) se la funzione parziale $f(\cdot, y_0)$ è derivabile in x_0 , ovvero se esiste finito il limite

$$\lim_{h \rightarrow 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h} \quad (3.14)$$

che denoteremo con uno dei simboli

$$\frac{\partial f}{\partial x}(x_0, y_0), \quad f_x(x_0, y_0), \quad D_x f(x_0, y_0).$$

A volte sottintenderemo il punto in cui si deriva scrivendo semplicemente

$$\frac{\partial f}{\partial x}, \quad f_x.$$

In modo analogo diremo che la f ammette derivata parziale rispetto alla seconda variabile se la funzione parziale $f(x_0, \cdot)$ è derivabile in y_0 , ovvero se esiste finito il limite

$$\lim_{k \rightarrow 0} \frac{f(x_0, y_0 + k) - f(x_0, y_0)}{k}$$

che denoteremo con uno dei simboli

$$\frac{\partial f}{\partial y}(x_0, y_0), \quad f_y(x_0, y_0), \quad D_y f(x_0, y_0).$$

Naturalmente, quanto detto si estende a funzioni definite su aperti di \mathbb{R}^N considerando le funzioni parziali rispetto a ciascuna variabile. Se $x = (x_1, \dots, x_N) \in R^N$ La f ammette derivata parziale i -esima per $i = 1, \dots, N$ nel punto x se

$$\frac{\partial f}{\partial x_i}(x) = \lim_{h \rightarrow 0} \frac{f(x_1, \dots, x_i + h, \dots, x_N) - f(x_1, \dots, x_N)}{h} = \lim_{h \rightarrow 0} \frac{f(x + he_i) - f(x)}{h} \in \mathbb{R}.$$

Se una funzione ammette derivate parziali allora chiaramente le funzioni parziali sono continue. Tuttavia è bene osservare che questo non implica necessariamente che la funzione stessa sia continua.

Esempio 121. *Si consideri la funzione definita in \mathbb{R}^2 nel modo seguente:*

$$f(x, y) = \begin{cases} \frac{xy}{x^2 + y^2} & (x, y) \neq (0, 0), \\ 0 & \text{altrimenti.} \end{cases}$$

Osserviamo che

$$\frac{\partial f}{\partial x}(0, 0) = \lim_{h \rightarrow 0} \frac{f(0 + h, 0) - f(0, 0)}{h} = 0,$$

poichè il numeratore è identicamente nullo. Allo stesso modo si trova che esiste la derivata parziale rispetto alla seconda variabile nell'origine e inoltre $f_y(0, 0) = 0$. Tuttavia osserviamo che

$$\lim_{x \rightarrow 0} f(x, x) = \lim_{x \rightarrow 0} \frac{x^2}{2x^2} = \frac{1}{2}.$$

Essendo $f(0, 0) = 0$, allora la funzione f non può essere continua nell'origine.

L'esistenza delle derivate parziali dunque implica soltanto la continuità delle funzioni parziali, ovvero della funzione lungo gli assi coordinati. In effetti, l'esempio precedente ha mostrato una funzione che pur essendo derivabile parzialmente non è continua ad esempio lungo la direzione della bisettrice $y = x$. In effetti, se una funzione è continua in un punto, i suoi valori si devono avvicinare qualunque sia il modo in cui ci si avvicina al punto. Pertanto, per dimostrare che una funzione non è continua spesso è vantaggioso cercare una curva o una successione per cui questo non avviene.

Esempio 122. Dire se la funzione

$$f(x, y) = \begin{cases} \frac{x^2 y}{x^4 + y^2} & (x, y) \neq (0, 0), \\ 0 & \text{altrimenti} \end{cases}$$

è continua. Nei punti diversi dall'origine la funzione è senz'altro continua essendo il rapporto di due polinomi il cui denominatore non si annulla mai. Resta allora da studiare il comportamento nell'origine. Se consideriamo ad esempio la parabola $y = x^2$ abbiamo che

$$f(x, x^2) = \frac{x^4}{2x^4} = \frac{1}{2}.$$

Essendo $f(0, 0) = 0$, la funzione non è continua nell'origine.

Come si intuisce da questi esempi, non esistono metodi che vadano bene per ogni situazione per determinare la continuità o, se vogliamo, per determinare i limiti di funzioni di più variabili e ogni caso deve essere studiato a parte.

Esempio 123. Dire se è continua la funzione

$$f(x, y) = \begin{cases} \frac{x^2 y}{x^2 + y^2} & (x, y) \neq (0, 0), \\ 0 & \text{altrimenti} \end{cases}$$

Evidentemente, è sufficiente esaminare il comportamento della funzione nell'origine. Si ha

$$|f(x, y) - f(0, 0)| = |f(x, y)| = \frac{x^2 |y|}{x^2 + y^2} \leq |y| \leq \|(x, y)\|.$$

Pertanto la funzione $f(x, y)$ è continua nell'origine e quindi in tutto \mathbb{R}^2 .

Definizione 124 (Gradiente). Sia $f : A \rightarrow \mathbb{R}$ definita su un aperto di \mathbb{R}^N . Supponiamo che f ammetta derivate parziali rispetto a tutte le variabili in un punto $x_0 \in A$. Allora si dice gradiente della f in x_0 il vettore

$$\text{grad}(f)(x_0) = \left(\frac{\partial f}{\partial x_1}(x_0), \dots, \frac{\partial f}{\partial x_N}(x_0) \right) \in \mathbb{R}^N.$$

Spesso useremo la notazione $\nabla f(x_0) := \text{grad}(f)(x_0)$.

Se $f : A \subset \mathbb{R}^N \rightarrow \mathbb{R}^M$, è utile introdurre le funzioni componenti mediante

$$f(x) = (f_1(x), \dots, f_M(x)) \in \mathbb{R}^M.$$

Per tali funzioni vettoriali diremo che la f è continua, o derivabile parzialmente, se lo sono tutte le funzioni componenti.

Definizione 125 (Matrice Jacobiana). Sia $f : A \subset \mathbb{R}^N \rightarrow \mathbb{R}^M$ derivabile parzialmente rispetto a tutte le variabili in un punto $x_0 \in A$. Si definisce matrice Jacobiana (o semplicemente Jacobiano) associata ad f la seguente matrice

$$Jf(x_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x_0) & \dots & \frac{\partial f_1}{\partial x_N}(x_0) \\ \frac{\partial f_2}{\partial x_1}(x_0) & \dots & \frac{\partial f_2}{\partial x_N}(x_0) \\ \dots & \dots & \dots \\ \frac{\partial f_M}{\partial x_1}(x_0) & \dots & \frac{\partial f_M}{\partial x_N}(x_0) \end{pmatrix} \in \mathcal{M}(M, N).$$

Definizione 126 (Derivate direzionali). Sia $A \subset \mathbb{R}^2$ aperto e sia $f : A \rightarrow \mathbb{R}$. Se $x_0 \in A$, esiste un $\delta > 0$ tale che $B(x_0, \delta) \subset A$. Fissato un qualsiasi vettore $v \in \mathbb{R}^N$ abbiamo che

$$\|x_0 - (x_0 + tv)\| = |t|\|v\| \Rightarrow x_0 + tv \in B(x_0, \delta), \quad \forall |t| < \frac{\delta}{\|v\|} := \rho.$$

Allora ha senso considerare la funzione φ definita da

$$t \in]-\rho, \rho[\mapsto f(x_0 + tv) \in \mathbb{R}.$$

Diremo allora che la funzione f ammette derivata direzionale lungo la direzione v nel punto x_0 , che indicheremo con il simbolo $\frac{\partial f}{\partial v}(x_0)$, se la funzione φ appena definita è derivabile nell'origine, ovvero se

$$\frac{\partial f}{\partial v}(x_0) := \lim_{t \rightarrow 0} \frac{f(x_0 + tv) - f(x_0)}{t} \in \mathbb{R}.$$

Pertanto le derivate parziali non sono altro che un caso particolare di derivate direzionali e precisamente le derivate lungo le direzioni degli assi coordinati.

3.4.2 Derivate successive e Teorema di Schwarz

Sia $A \subset \mathbb{R}^2$ aperto e sia $f : A \rightarrow \mathbb{R}$ dotata di derivate parziali in un punto (x_0, y_0) . Se la funzione

$$f_x : A \rightarrow \mathbb{R}$$

è a sua volta dotata di derivate parziale rispetto ad x diremo che la f è derivabile due volte rispetto ad x nel punto (x_0, y_0) e denoteremo tale derivata con uno dei simboli

$$D_{xx}f(x_0, y_0), \quad f_{xx}(x_0, y_0), \quad \frac{\partial^2 f}{\partial x^2}(x_0, y_0).$$

Se invece f_x è derivabile rispetto alla y diremo che la f è derivabile due volte rispetto ad x e a y e indicheremo la derivata con

$$D_{xy}f(x_0, y_0), \quad f_{xy}(x_0, y_0), \quad \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0).$$

Analogamente, se f_y ammette derivate parziali potremo considerare le derivate seconde f_{yy} e f_{yx} . Talvolta le f_{xx} e f_{yy} vengono dette derivate seconde pure, mentre le f_{xy} e f_{yx} vengono dette derivate seconde miste. Allo stesso modo si considerano le derivate di ordine maggiore di due. Nel caso delle funzioni a valori vettoriali si dovranno considerare naturalmente le derivate delle funzioni componenti.

Definizione 127. Sia $A \subset \mathbb{R}^2$ aperto e sia $f : A \rightarrow \mathbb{R}$. Se $k \geq 1$ è un intero, diremo che $f \in \mathcal{C}^k(A)$, o semplicemente che f è di classe \mathcal{C}^k in A , se la f ammette le derivate parziali fino all'ordine k e queste ultime sono continue.

In generale le derivate seconde dipendono dall'ordine in cui si esegue la derivazione. Ad esempio, se consideriamo la funzione

$$f(x, y) = \begin{cases} y^2 \arctan \frac{x}{y} & y \neq 0, \\ 0 & \text{altrimenti} \end{cases}$$

il lettore può verificare per esercizio che risulta

$$\frac{\partial^2 f}{\partial x \partial y}(0, 0) = 0, \quad \frac{\partial^2 f}{\partial y \partial x}(0, 0) = 1.$$

Il teorema seguente fornisce una condizione sufficiente a garantire che questo non accada, cioè che si possa invertire l'ordine di derivazione.

Teorema 128 (di Schwarz). *Sia $A \subset \mathbb{R}^2$ ed $f \in \mathcal{C}^2(A)$. Allora $f_{xy} = f_{yx}$.*

Dimostrazione. Fissiamo un punto $(x_0, y_0) \in A$. Dobbiamo verificare che $f_{xy}(x_0, y_0) = f_{yx}(x_0, y_0)$. Poichè A è aperto esiste un intervallo aperto $I \subset \mathbb{R}$ tale che $(x_0, y_0) \in I \times I \subset A$. Allora possiamo definire le funzioni

$$\begin{aligned} F(x) &= f(x, y) - f(x, y_0), & x \in I, \\ G(y) &= f(x, y) - f(x_0, y), & y \in I. \end{aligned}$$

Si verifica immediatamente che

$$F(x) - F(x_0) = G(y) - G(y_0) \quad \forall (x, y) \in I \times I. \quad (3.15)$$

Per il Teorema di Lagrange applicato alla funzione F nell'intervallo $[x_0, x]$, esiste un $x' \in [x_0, x]$ tale che

$$F(x) - F(x_0) = F'(x')(x - x_0) = [f_x(x', y) - f_x(x', y_0)](x - x_0).$$

Applicando di nuovo il Teorema di Lagrange alla funzione $f_x(x', \cdot)$ si trova un $y' \in [y_0, y]$ tale che

$$F(x) - F(x_0) = [f_x(x', y) - f_x(x', y_0)](x - x_0) = f_{yx}(x', y')(y - y_0)(x - x_0). \quad (3.16)$$

Applicando in modo analogo il Teorema di Lagrange alla funzione G troviamo $x'' \in [x_0, x]$ e $y'' \in [y_0, y]$ tali che

$$G(y) - G(y_0) = f_{xy}(x'', y'')(y - y_0)(x - x_0). \quad (3.17)$$

Mettendo insieme (3.15), (3.16) e (3.17) otteniamo

$$f_{yx}(x', y') = f_{xy}(x'', y'').$$

Poiché le derivate seconde sono continue, passando al limite per $x \rightarrow x_0$ e $y \rightarrow y_0$ si ottiene finalmente che

$$f_{yx}(x_0, y_0) = f_{xy}(x_0, y_0). \quad \square$$

Il Teorema di Schwarz vale naturalmente anche per funzioni definite su aperti di \mathbb{R}^N .

Definizione 129 (Matrice Hessiana). *Sia $f : A \rightarrow \mathbb{R}$ una funzione definita su di un aperto di \mathbb{R}^N e dotata delle derivate parziali seconde. Allora si definisce la matrice Hessiana nel punto $x_0 \in A$ come la matrice*

$$H_f(x_0) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(x_0) & \frac{\partial^2 f}{\partial x_1 x_2}(x_0) & \dots & \frac{\partial^2 f}{\partial x_1 x_N}(x_0) \\ \frac{\partial^2 f}{\partial x_2 x_1}(x_0) & \frac{\partial^2 f}{\partial x_2^2}(x_0) & \dots & \frac{\partial^2 f}{\partial x_2 x_N}(x_0) \\ \dots & \dots & \dots & \dots \\ \frac{\partial^2 f}{\partial x_N x_1}(x_0) & \frac{\partial^2 f}{\partial x_N x_2}(x_0) & \dots & \frac{\partial^2 f}{\partial x_N^2}(x_0) \end{pmatrix} \in \mathcal{M}(N, N).$$

Se f è di classe \mathcal{C}^2 , allora per il Teorema di Schwarz la matrice H_f è simmetrica.

3.4.3 Differenziale di funzioni di più variabili

Il concetto che permette di estendere a \mathbb{R}^N molte delle proprietà delle funzioni di una variabile derivabili è quello più restrittivo di differenziale.

Definizione 130 (Differenziale). *Sia $A \subset \mathbb{R}^N$ aperto e $f : A \rightarrow \mathbb{R}^M$. La funzione f si dice differenziabile in un punto $x_0 \in A$ se esiste un'applicazione lineare $L : \mathbb{R}^N \rightarrow \mathbb{R}^M$ tale che*

$$\lim_{\|h\| \rightarrow 0} \frac{f(x_0 + h) - f(x_0) - L(h)}{\|h\|} = 0. \quad (3.18)$$

L'applicazione L si dice differenziale della f in x_0 e si indica con il simbolo

$$df(x_0).$$

Definizione 131 (Forme quadratiche). Sia $A \in \mathcal{M}(N)$ una matrice simmetrica. Si dice forma quadratica associata ad A la funzione $q_A : \mathbb{R}^N \rightarrow \mathbb{R}$ definita da

$$q_A(x) = \langle Ax, x \rangle.$$

La matrice A (o la forma quadratica q_A associata) si dice definita positiva se esiste $m > 0$ tale che

$$\forall x \in \mathbb{R}^N, x \neq 0, \quad q_A(x) \geq m\|x\|^2 > 0.$$

Si dice definita negativa se esiste $M < 0$ tale che

$$\forall x \in \mathbb{R}^N, x \neq 0, \quad q_A(x) \leq M\|x\|^2 < 0.$$

La matrice A si dice positiva se

$$\forall x \in \mathbb{R}^N \quad q_A(x) \geq 0$$

e negativa se

$$\forall x \in \mathbb{R}^N \quad q_A(x) \leq 0.$$

Osserviamo che la funzione q_A è continua. Considerato allora l'insieme compatto

$$S = \{x \in \mathbb{R}^N : \|x\| = 1\},$$

la funzione q_A ammetterà un minimo m ed un massimo M su S . Essendo A simmetrica sappiamo che A ammette autovalori. Dunque esisteranno anche il più piccolo e il più grande degli autovalori. Sia λ ad esempio il più piccolo autovalore e sia v un autovettore corrispondente. Allora abbiamo

$$m \leq q_A \left(\frac{v}{\|v\|} \right) = \left\langle A \frac{v}{\|v\|}, \frac{v}{\|v\|} \right\rangle = \left\langle \lambda \frac{v}{\|v\|}, \frac{v}{\|v\|} \right\rangle = \lambda.$$

Bibliografia

- [1] E. Acerbi, G. Buttazzo, Primo Corso di Analisi Matematica, Pitagora Editrice, 1997.
- [2] J. P. Cecconi, G. Stampacchia, Analisi Matematica, 2 Volume, funzioni di più variabili, Liguori editore, 1983.
- [3] S. Lang, Algebra Lineare, Bollati Boringhieri, 1970.
- [4] P. Marcellini, C. Sbordone, Esercitazioni di Matematica, 1 Volume parte seconda, Liguori Editore, 1989.
- [5] P. Marcellini, C. Sbordone, Esercitazioni di Matematica, 2 Volume parte prima, Liguori Editore, 1989.
- [6] P. Marcellini, C. Sbordone, Esercitazioni di Matematica, 2 Volume parte seconda, Liguori Editore, 1991.
- [7] E. Giusti, Esercizi e Complementi di Analisi Matematica, 1 Volume, Bollati Boringhieri, 2000.
- [8] E. Giusti, Esercizi e Complementi di Analisi Matematica, 2 Volume, Bollati Boringhieri, 2000.