
Fabrizio Bonani

Matematica III

LAUREA DI I LIVELLO A DISTANZA
IN INGEGNERIA INFORMATICA
POLO TECNOLOGICO DI ALESSANDRIA
ANNO ACCADEMICO 2000/2001

Indice

1	Funzioni di più variabili reali	1
1.1	Richiami teorici: distanza, limite e derivate parziali	1
1.1.1	Elementi di topologia in \mathbb{R}^n	1
1.1.2	Funzioni, limiti e continuità	4
1.1.3	Calcolo differenziale per funzioni di più variabili	7
1.2	Estremi liberi	10
1.3	Estremi vincolati	18
2	Equazioni differenziali	23
2.1	Introduzione	23
2.2	Equazioni differenziali a variabili separabili	24
2.3	Equazioni differenziali lineari del primo ordine	27
2.4	Sistemi di equazioni differenziali lineari a coefficienti costanti	33
2.5	Equazioni differenziali scalari lineari a coefficienti costanti	39
3	Integrali multipli	45
3.1	Integrali nel piano	45
3.2	Integrali nello spazio	53
3.3	Cambiamenti di variabile	58
4	Integrali di linea e di flusso	63
4.1	Richiami su curve e superfici	63
4.1.1	Curve	63
4.1.2	Superfici	66
4.2	Integrali di linea	68
4.3	Integrali di flusso	75

5	Calcolo delle probabilità	79
5.1	Elementi di teoria della probabilità	79
5.2	Calcolo combinatorio	86
5.3	Variabili casuali discrete	92
5.3.1	Densità discreta e funzione di ripartizione	94
5.3.2	Speranza matematica	98
5.4	Variabili casuali continue	105
5.4.1	Funzioni di variabili casuali assolutamente continue	108
5.4.2	Variabili casuali gaussiane	111
5.4.3	Speranza matematica e momenti	115

1

Funzioni di più variabili reali

Obiettivi: Calcolo differenziale di funzioni di più variabili. Estremi liberi di una funzione di più variabili. Cenni sugli estremi vincolati.

1.1 Richiami teorici: distanza, limite e derivate parziali

1.1.1 Elementi di topologia in \mathbb{R}^n

Nell'ambito di questo capitolo prenderemo in esame funzioni definite su \mathbb{R}^n , ovvero sull'insieme di tutte le n -uple ordinate di numeri reali:

$$\mathbb{R}^n = \{(x_1, x_2, \dots, x_n) : x_i \in \mathbb{R} \forall i\}.$$

Particolarmente importanti nelle applicazioni che verranno considerate nel seguito sono il *piano* \mathbb{R}^2 e lo *spazio* \mathbb{R}^3 . Un elemento $\mathbf{x} \in \mathbb{R}^n$ rappresenta un *vettore* in \mathbb{R}^n di componenti $\mathbf{x} = (x_1, x_2, \dots, x_n)$.

Dati due elementi $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ e uno scalare $\lambda \in \mathbb{R}$, si definiscono le due operazioni:

- *somma* dei due elementi $\mathbf{z} = \mathbf{x} + \mathbf{y}$, caratterizzata da:

$$(z_1, \dots, z_n) = (x_1 + y_1, \dots, x_n + y_n);$$

- *prodotto per uno scalare* $\mathbf{z} = \lambda \mathbf{x}$, caratterizzato da:

$$(z_1, \dots, z_n) = (\lambda x_1, \dots, \lambda x_n).$$

Con *prodotto scalare* di due elementi $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ si intende invece:

$$\mathbf{x} \cdot \mathbf{y} = \sum_i x_i y_i.$$

Due vettori \mathbf{x}, \mathbf{y} si dicono *ortogonali* (risp. *paralleli*) se $\mathbf{x} \cdot \mathbf{y} = 0$ (risp. $\exists \lambda \in \mathbb{R}: \mathbf{x} = \lambda \mathbf{y}$).

Il prodotto scalare consente di definire il *modulo* di un vettore \mathbf{x} :

$$|\mathbf{x}| = \sqrt{\mathbf{x} \cdot \mathbf{x}} = \left(\sum_i x_i^2 \right)^{1/2},$$

che soddisfa le seguenti proprietà:

- positività: $|\mathbf{x}| \geq 0$ e $|\mathbf{x}| = 0 \iff \mathbf{x} = \mathbf{0}$;
- omogeneità: $|\lambda \mathbf{x}| = |\lambda| |\mathbf{x}|$;
- disuguaglianza triangolare: $|\mathbf{x} + \mathbf{y}| \leq |\mathbf{x}| + |\mathbf{y}|$;
- disuguaglianza di Cauchy-Schwartz: $|\mathbf{x} \cdot \mathbf{y}| \leq |\mathbf{x}| |\mathbf{y}|$.

Sulla base della definizione di modulo, si introduce il concetto di *distanza* d in \mathbb{R}^n :

$$d(\mathbf{x}, \mathbf{y}) = |\mathbf{x} - \mathbf{y}|,$$

che gode delle seguenti proprietà, direttamente derivabili da quelle del modulo:

- positività e annullamento: $d(\mathbf{x}, \mathbf{y}) \geq 0$ e $d(\mathbf{x}, \mathbf{y}) = 0 \iff \mathbf{x} = \mathbf{y}$;
- simmetria: $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$;
- disuguaglianza triangolare: $d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y})$.

Grazie a queste definizioni, si possono introdurre alcuni concetti topologici fondamentali:

Definizione 1.1.1 Si chiama *sfera* o *intorno* di centro \mathbf{x}_0 e raggio $r > 0$ l'insieme:

$$B(\mathbf{x}_0, r) = \{\mathbf{x} \in \mathbb{R}^n : d(\mathbf{x}, \mathbf{x}_0) < r\}.$$

Definizione 1.1.2 Un insieme $A \subset \mathbb{R}^n$ si dice *limitato* se scelto un qualunque punto $\mathbf{x}_0 \in A$, esiste una sfera $B(\mathbf{x}_0, r)$ di centro \mathbf{x}_0 e raggio $r > 0$ che lo contiene.

Definizione 1.1.3 Si consideri un insieme $A \subset \mathbb{R}^n$. Un punto $\mathbf{x} \in A$ si dice *interno* ad A se esiste un intorno $B(\mathbf{x}, r)$, con $r > 0$, contenuto in A . Un punto $\mathbf{x} \notin A$ si dice *esterno* ad A se è interno al suo complementare. Si chiama *frontiera* di A , indicata con ∂A , l'insieme dei punti di \mathbb{R}^n che non sono nè interni nè esterni ad A . Si chiama *chiusura* di A , indicata con \overline{A} , l'unione di A con la sua frontiera, mentre si chiama *parte interna* di A , indicata con $\overset{\circ}{A}$, l'insieme dei punti interni ad A .

Definizione 1.1.4 Un insieme $A \subset \mathbb{R}^n$ si dice *aperto* se scelto un qualunque punto $\mathbf{x}_0 \in A$, esiste una sfera $B(\mathbf{x}_0, r)$ di centro \mathbf{x}_0 e raggio $r > 0$ contenuta in A .

Definizione 1.1.5 Un insieme $C \subset \mathbb{R}^n$ si dice *chiuso* se il suo complementare $\mathbb{R}^n \setminus C$ è aperto.

Una applicazione diretta di queste definizioni consente di verificare come, in generale, gli insiemi aperti siano caratterizzati dall'essere regioni di \mathbb{R}^n esclusa la frontiera. Ad esempio, l'insieme $\{\mathbf{x} \in \mathbb{R}^n : d(\mathbf{x}, \mathbf{0}) < 1\} = B(\mathbf{0}, 1)$ è aperto, mentre $\{\mathbf{x} \in \mathbb{R}^n : d(\mathbf{x}, \mathbf{0}) \leq 1\}$ è chiuso. Naturalmente, possono esistere insiemi che non sono nè chiusi nè aperti, ad esempio:

$$\{\mathbf{x} \in \mathbb{R}^2 : 0 \leq x_1 \leq 1 \quad \& \quad 0 < x_2 < 1\}.$$

Si può inoltre dimostrare:

Teorema 1.1.1 L'unione e l'intersezione di un numero finito di insiemi aperti (risp. chiusi) sono insiemi aperti (risp. chiusi).

Teorema 1.1.2 Sia $A \subset \mathbb{R}^n$. Allora:

1. ∂A coincide con la frontiera di $\mathbb{R}^n \setminus A$;
2. A è chiuso se e solo se $A = \overline{A}$;
3. A è aperto se e solo se $\overset{\circ}{A}$;
4. $\overset{\circ}{A} = A \setminus \partial A$.

Definizione 1.1.6 Un insieme $A \subset \mathbb{R}^n$ si dice *compatto* se è chiuso, limitato e non vuoto.

Nel seguito del Capitolo, a meno di una definizione esplicita, considereremo sempre insiemi aperti.

1.1.2 Funzioni, limiti e continuità

Una *funzione reale di più variabili*, detta anche *campo scalare*, è una applicazione $f: A \subset \mathbb{R}^n \rightarrow \mathbb{R}$, mentre con *funzione vettoriale* si intende una applicazione $\mathbf{F}: A \subset \mathbb{R}^m \rightarrow \mathbb{R}^n$. Nel caso in cui $m = n$, la funzione vettoriale viene anche detta *campo vettoriale*. In particolare, una *curva* in \mathbb{R}^n è una funzione vettoriale $\Phi: I \subset \mathbb{R} \rightarrow \mathbb{R}^n$ definita su un intervallo I ; se I è un intervallo chiuso e limitato, si parla di *arco di curva*. Infine, una *superficie* in \mathbb{R}^3 è una funzione vettoriale definita in \mathbb{R}^2 , ovvero $\Sigma: A \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$.

Si consideri una successione di punti $\mathbf{a}_k = (a_1^{(k)}, \dots, a_n^{(k)}) \in \mathbb{R}^n$.

Definizione 1.1.7 La successione $\{\mathbf{a}_k\}$ di punti di \mathbb{R}^n converge al punto $\mathbf{x}_0 \in \mathbb{R}^n$ se, dato un qualunque $\epsilon > 0$, esiste un intero $k(\epsilon)$ tale che per ogni $n \geq k(\epsilon)$:

$$|\mathbf{a}_k - \mathbf{x}_0| < \epsilon \iff \mathbf{a}_k \in B(\mathbf{x}_0, \epsilon).$$

In questo caso, si scrive:

$$\lim_{k \rightarrow \infty} \mathbf{a}_k = \mathbf{x}_0.$$

Si può dimostrare:

Teorema 1.1.3 La successione $\{\mathbf{a}_k\} \subset \mathbb{R}^n$ converge ad $\mathbf{x}_0 \in \mathbb{R}^n$ se e solo se, per ogni $1 \leq i \leq n$:

$$\lim_{k \rightarrow \infty} a_i^{(k)} = x_0^{(i)}.$$

Il concetto di limite viene esteso alle funzioni vettoriali $\mathbf{F}: A \subset \mathbb{R}^m \rightarrow \mathbb{R}^n$ grazie a:

Definizione 1.1.8 La funzione vettoriale \mathbf{F} converge a $\mathbf{l} \in \mathbb{R}^n$ nel punto $\mathbf{x}_0 \in \mathbb{R}^m$ se per ogni successione $\{\mathbf{a}_k\} \subset \mathbb{R}^m$ convergente a \mathbf{x}_0 , la successione $\{\mathbf{F}(\mathbf{a}_k)\} \subset \mathbb{R}^n$ converge a \mathbf{l} . In questo caso, si scrive:

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \mathbf{F}(\mathbf{x}) = \mathbf{l}.$$

Dal Teorema 1.1.3, segue immediatamente che anche per le funzioni vettoriali il limite viene definito componente per componente.

Teorema 1.1.4 Si considerino le funzioni vettoriali $\mathbf{F}, \mathbf{F}_1, \mathbf{F}_2: A \subset \mathbb{R}^m \rightarrow B \subset \mathbb{R}^n$, $\mathbf{G}: B \subset \mathbb{R}^n \rightarrow \mathbb{R}^p$ e $\mathbf{F}_3: A \subset \mathbb{R}^m \rightarrow \mathbb{R}^p$.

1. Se esiste il limite di $\mathbf{F}(\mathbf{x})$ per $\mathbf{x} \rightarrow \mathbf{x}_0$, allora questo è unico (*unicità del limite*).

2. Se $\mathbf{F}_1(\mathbf{x})$ e $\mathbf{F}_2(\mathbf{x})$ convergono, rispettivamente, a \mathbf{l}_1 e \mathbf{l}_2 per $\mathbf{x} \rightarrow \mathbf{x}_0$, allora $\mathbf{F}_1(\mathbf{x}) + \mathbf{F}_2(\mathbf{x})$ converge a $\mathbf{l}_1 + \mathbf{l}_2$ per $\mathbf{x} \rightarrow \mathbf{x}_0$. Inoltre, se $\mathbf{F}(\mathbf{x})$ converge a \mathbf{l} per $\mathbf{x} \rightarrow \mathbf{x}_0$, allora $\lambda \mathbf{F}(\mathbf{x})$ ($\lambda \in \mathbb{R}$) converge a $\lambda \mathbf{l}$ per $\mathbf{x} \rightarrow \mathbf{x}_0$ (*limite della somma e del prodotto per una costante*).
3. Si consideri la funzione composta $\mathbf{F}_3 = \mathbf{G} \circ \mathbf{F}$, dove $\mathbf{F}(\mathbf{x})$ converge a \mathbf{l} per $\mathbf{x} \rightarrow \mathbf{x}_0$ e \mathbf{l} è nel dominio di \mathbf{G} . Se $\mathbf{G}(\mathbf{x})$ converge a \mathbf{l}_1 per $\mathbf{x} \rightarrow \mathbf{l}$, allora $\mathbf{F}_3(\mathbf{x})$ converge a \mathbf{l}_1 per $\mathbf{x} \rightarrow \mathbf{x}_0$ (*limite della funzione composta*).

Teorema 1.1.5 Si considerino i campi scalari $f, g: A \subset \mathbb{R}^n \rightarrow \mathbb{R}$.

1. Se $f(\mathbf{x})$ e $g(\mathbf{x})$ convergono, rispettivamente, a l_1 e l_2 per $\mathbf{x} \rightarrow \mathbf{x}_0$, allora $f(\mathbf{x})g(\mathbf{x})$ converge a $l_1 l_2$ e, se $g(\mathbf{x}), l_2 \neq 0$, $f(\mathbf{x})/g(\mathbf{x})$ converge a l_1/l_2 (*limite del prodotto e del quoziente*).
2. Se $f(\mathbf{x})$ converge a l per $\mathbf{x} \rightarrow \mathbf{x}_0$ ed esiste una sfera di centro \mathbf{x}_0 e raggio r nella quale $f(\mathbf{x}) \geq 0$, allora anche $l \geq 0$ (*permanenza del segno*).

La continuità di una funzione di più variabili viene definita in analogia a quanto avviene nel caso delle funzioni di una sola variabile:

Definizione 1.1.9 La funzione vettoriale $\mathbf{F}: A \subset \mathbb{R}^m \rightarrow \mathbb{R}^n$ è *continua* in $\mathbf{x}_0 \in A$ se:

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \mathbf{F}(\mathbf{x}) = \mathbf{F}(\mathbf{x}_0).$$

Evidentemente, grazie al Teorema 1.1.3 una funzione vettoriale è continua se e solo se lo sono tutte le sue componenti. Inoltre, dai Teoremi 1.1.4 e 1.1.5 segue che somma, prodotto, quoziente e prodotto di composizione (quando queste operazioni hanno senso) di funzioni continue sono a loro volta funzioni continue. In particolare, una funzione continua di una variabile, è continua anche se considerata come funzione di più variabili.

Esempio 1.1.1

Le funzioni di due variabili:

$$f(x,y) = \cos x \quad g(x,y) = \cos y$$

sono continue in quanto la funzione $h: \mathbb{R} \rightarrow \mathbb{R}$ $h(t) = \cos t$ è continua. ◆

Esempio 1.1.2

La funzione:

$$\mathbf{F}(x,y) = (f_1(x,y), f_2(x,y)) = \left(x^3 e^{x+y}, \frac{\sin y}{1+x^2} \right)$$

è continua in \mathbb{R}^2 . Infatti:

- la prima componente $f_1(x,y)$ è costituita dal prodotto di x^3 (funzione continua in \mathbb{R} , e quindi anche in \mathbb{R}^2) e di una funzione composta a sua volta continua. Infatti, x e y sono funzioni continue in \mathbb{R}^2 , e quindi anche $x+y$. Infine, l'esponenziale è una funzione continua in \mathbb{R} , quindi anche in \mathbb{R}^2 , e per la continuità del prodotto di composizione segue la tesi;
- la seconda componente $f_2(x,y)$ è costituita dal rapporto di due funzioni continue: $\sin y$ e $1+x^2$, infatti entrambe sono funzioni continue in \mathbb{R} , e quindi anche in \mathbb{R}^2 .



Per le funzioni continue, vale l'importante:

Teorema 1.1.6 (di Weierstrass) Sia $f: A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ una funzione continua sull'insieme compatto A . Allora f ammette massimo e minimo in A .

Qualora si voglia dimostrare il fatto che una funzione $\mathbf{F}(\mathbf{x})$ non ammetta limite in \mathbf{x}_0 , una tecnica efficace consiste nel valutare tale limite su una restrizione della funzione ad una curva passante per il punto \mathbf{x}_0 : se si riescono a trovare due curve per le quali la restrizione tende a valori diversi, per il Teorema di unicità del limite si è provata la non convergenza nel punto in esame.

Esempio 1.1.3

Calcolare:

$$\lim_{\mathbf{x} \rightarrow \mathbf{0}} f(\mathbf{x}) = \lim_{\mathbf{x} \rightarrow \mathbf{0}} \frac{xy}{x^2 + y^2} \quad \mathbf{x} \in \mathbb{R}^2.$$

Consideriamo la retta passante per l'origine $y = x$: la restrizione della funzione $f(x,y)$ su questa retta vale:

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

pertanto lungo tale restrizione il limite vale $1/2$.

Un'altra retta passante per l'origine del piano è $y = -x$. La restrizione di f a questa curva è:

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

pertanto il limite calcolato lungo questa direzione vale $-1/2$. Per il teorema di unicità del limite, quindi, si può concludere che il limite cercato non esiste. ♦

1.1.3 Calcolo differenziale per funzioni di più variabili

Nell'ambito di questa sezione, prenderemo in considerazione unicamente funzioni reali di più variabili, ovvero campi scalari $f: A \subset \mathbb{R}^n \rightarrow \mathbb{R}$. Dato un *versore* $\hat{\mathbf{v}} \in \mathbb{R}^n$, ovvero un vettore di lunghezza unitaria:

$$|\hat{\mathbf{v}}| = 1,$$

si definisce:

Definizione 1.1.10 La *derivata direzionale* della funzione f nella direzione definita dal versore $\hat{\mathbf{v}}$ nel punto \mathbf{x} è la quantità:

$$D_{\hat{\mathbf{v}}}f(\mathbf{x}) = \frac{df}{d\hat{\mathbf{v}}}(\mathbf{x}) = f_{\hat{\mathbf{v}}}(\mathbf{x}) = \lim_{\theta \rightarrow 0} \frac{f(\mathbf{x} + \theta\hat{\mathbf{v}}) - f(\mathbf{x})}{\theta}.$$

In particolare, se il versore è diretto lungo uno degli assi dello spazio \mathbb{R}^n , si ha:

Definizione 1.1.11 La *derivata parziale* della funzione f rispetto alla variabile x_i nel punto \mathbf{x} è la quantità:

$$\frac{\partial f}{\partial x_i}(\mathbf{x}) = f_{x_i}(\mathbf{x}) = \lim_{\theta \rightarrow 0} \frac{f(x_1, \dots, x_i + \theta, \dots, x_n) - f(x_1, \dots, x_i, \dots, x_n)}{\theta}.$$

Sulla base di questa definizione, segue immediatamente che una funzione di n variabili reali è dotata al più di n derivate parziali, che possono essere singolarmente valutate *applicando le consuete regole di derivazione di funzioni di una sola variabile, pur di considerare tutte le componenti di \mathbf{x} diverse da x_i come costanti*.

Esempio 1.1.4

Calcolare tutte le derivate parziali della funzione di due variabili:

$$f(\mathbf{x}) = f(x, y) = e^{x^2 + y^2} \sin(x + y).$$

Iniziamo a calcolare la derivata parziale rispetto alla variabile x :

$$\frac{\partial f}{\partial x} = 2xe^{x^2 + y^2} \sin(x + y) + e^{x^2 + y^2} \cos(x + y).$$

Per la derivata parziale rispetto a y si procede analogamente:

$$\frac{\partial f}{\partial y} = 2ye^{x^2 + y^2} \sin(x + y) + e^{x^2 + y^2} \cos(x + y).$$



Di particolare importanza è la funzione vettoriale avente per componenti le derivate parziali di un campo scalare: essa viene detta *gradiente* di f , e viene di solito indicata con la notazione ∇f .

Definizione 1.1.12 Una funzione reale di più variabili è detta *derivabile* in un punto \mathbf{x} del suo dominio se in esso esistono tutte le derivate parziali della funzione stessa.

Si noti come, in perfetta analogia al caso delle funzioni di variabile reale, si possano definire *derivate parziali di ordine p* . In particolare, le derivate parziali seconde di una funzione di n variabili reali sono n^2 :

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = f_{x_i x_j} \quad i, j = 1, \dots, n,$$

dove naturalmente:

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_j} \right).$$

In realtà, delle n^2 derivate parziali seconde solo $n(n+1)/2$ sono indipendenti, purché la funzione sia sufficientemente regolare:

Teorema 1.1.7 (di Schwarz) Se le derivate parziali seconde $f_{x_i x_j}$, $f_{x_j x_i}$ sono entrambe definite in un aperto A e continue in $\mathbf{x}_0 \in A$, allora in tale punto coincidono.

In pratica, le derivate parziali di ordine superiore al primo, se la funzione è sufficientemente regolare, sono indipendenti dall'ordine di derivazione. Infatti, il Teorema 1.1.7 può essere generalizzato a derivate parziali di ordine qualsiasi.

Contrariamente a quanto avviene in \mathbb{R} , per le funzioni di più variabili *derivabilità* (cioè l'esistenza dei limiti dei rapporti incrementali che ne definiscono le derivate

parziali) e *differenziabilità* (cioè l'esistenza del *piano tangente*) non coincidono. In particolare, con *piano tangente* ad un campo scalare f in un punto $\mathbf{x}_0 \in \text{dom}\{f\}$ si intende una applicazione da \mathbb{R}^n in \mathbb{R} lineare in $(\mathbf{x} - \mathbf{x}_0)$ equivalente a f per $\mathbf{x} \rightarrow \mathbf{x}_0$, ovvero:

$$f(\mathbf{x}) - f(\mathbf{x}_0) = \mathbf{F}(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) + o(|\mathbf{x} - \mathbf{x}_0|). \quad (1.1a)$$

Si può dimostrare che, se il vettore $\mathbf{F}(\mathbf{x}_0)$ che soddisfa la (1.1a) esiste, esso coincide con il gradiente di f calcolato in \mathbf{x}_0 :

$$\mathbf{F}(\mathbf{x}_0) = \nabla f(\mathbf{x}_0). \quad (1.1b)$$

Il campo scalare $\nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0)$ definisce il *piano tangente* al campo scalare nel punto \mathbf{x}_0 .

Definizione 1.1.13 Il campo scalare f è detto *differenziabile* in \mathbf{x}_0 se la (1.1) è verificata.

Teorema 1.1.8 Sia $f: A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ un campo scalare. Se tutte le derivate parziali di f esistono e sono continue in A , allora f è differenziabile in A .

Una funzione che soddisfi le ipotesi di questo teorema è detta di *classe* $C^1(A)$, e si scrive $f \in C^1(A)$.

Il gradiente di una funzione di più variabili è direttamente correlato alla derivata direzionale, poiché si può dimostrare:

Teorema 1.1.9 Sia $f: A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ un campo scalare differenziabile nel punto $\mathbf{x}_0 \in A$. Allora, per ogni versore $\hat{\mathbf{v}}$ esiste la derivata direzionale $D_{\hat{\mathbf{v}}}f(\mathbf{x}_0)$ nel punto \mathbf{x}_0 , e vale:

$$D_{\hat{\mathbf{v}}}f(\mathbf{x}_0) = \nabla f(\mathbf{x}_0) \cdot \hat{\mathbf{v}}.$$

Di conseguenza, una qualsiasi derivata direzionale per una funzione differenziabile può essere espressa come combinazione lineare delle derivate parziali della funzione stessa.

Le regole per il calcolo del gradiente possono essere direttamente derivate dalle proprietà delle derivate parziali. In particolare, ricordando che con *curva regolare* si intende una curva $\Phi(t) : I \subset \mathbb{R} \rightarrow \mathbb{R}^n$ dotata di vettore tangente:

$$\frac{d\Phi}{dt} = \Phi'(t)$$

non nullo per ogni t , si possono dimostrare:

Teorema 1.1.10 Siano f e g funzioni reali derivabili di più variabili. Allora valgono le seguenti proprietà del gradiente:

$$\begin{aligned}\nabla(\alpha f + \beta g) &= \alpha \nabla f + \beta \nabla g & \alpha, \beta \in \mathbb{R} \\ \nabla(fg) &= g \nabla f + f \nabla g.\end{aligned}$$

Teorema 1.1.11 (Derivazione delle funzioni composte) Sia $f: A \subset \mathbb{R}^n \rightarrow B \subset \mathbb{R}$ differenziabile, e siano $g: B \subset \mathbb{R} \rightarrow \mathbb{R}$ una funzione derivabile e $\Phi: I \subset \mathbb{R} \rightarrow A \subset \mathbb{R}^n$ un arco di curva regolare. Posto:

$$h_1(\mathbf{x}) = g(f(\mathbf{x})) \qquad h_2(t) = f(\Phi(t))$$

allora:

$$\begin{aligned}\nabla h_1(\mathbf{x}) &= g'(f(\mathbf{x})) \nabla f(\mathbf{x}) \\ h_2'(t) &= \nabla f(\Phi(t)) \cdot \Phi'(t).\end{aligned}$$

1.2 Estremi liberi

Gli strumenti del calcolo differenziale per funzioni di più variabili introdotti nella Sez. 1.1 possono essere utilizzati nello studio dei loro estremi, ovvero:

Definizione 1.2.1 Sia $f: A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ e sia $\mathbf{x}_0 \in A$. Si dice che \mathbf{x}_0 è un punto di *massimo* (risp. *minimo*) *assoluto* per f in A se $f(\mathbf{x}) \leq f(\mathbf{x}_0) \forall \mathbf{x} \in A$ (risp. $f(\mathbf{x}) \geq f(\mathbf{x}_0) \forall \mathbf{x} \in A$). Si dice invece che \mathbf{x}_0 è un punto di *massimo* (risp. *minimo*) *relativo* o *locale* se esiste un intorno $B(\mathbf{x}_0, r)$ di \mathbf{x}_0 tale che $f(\mathbf{x}) \leq f(\mathbf{x}_0) \forall \mathbf{x} \in B(\mathbf{x}_0, r)$ (risp. $f(\mathbf{x}) \geq f(\mathbf{x}_0) \forall \mathbf{x} \in B(\mathbf{x}_0, r)$).

In questa sezione, ci occuperemo della ricerca dei punti di minimo e massimo nell'intero dominio di definizione della funzione considerata. Il caso degli *estremi vincolati*, ovvero la ricerca dei punti di massimo e minimo per una restrizione di una funzione di più variabili, verrà brevemente trattato nella Sez. 1.3.

Di importanza fondamentale nella ricerca degli estremi di una funzione di più variabili è il seguente

Teorema 1.2.1 (di Fermat) Sia $f: A \subset \mathbb{R}^n \rightarrow \mathbb{R}$, e sia $\mathbf{x}_0 \in A$ un punto di massimo o minimo locale per f nel quale la funzione sia derivabile. Allora, $\nabla f(\mathbf{x}_0) = \mathbf{0}$.

Definizione 1.2.2 I punti $\mathbf{x} \in A$ in cui il gradiente di una funzione $f: A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ si annulla sono detti *punti critici* o *stazionari* di f .

Esempio 1.2.1

Si cerchino i punti di massimo e minimo della funzione:

$$f(\mathbf{x}) = f(x,y) = 3x^2 + y^2 - x^3y.$$

Il dominio della funzione f è evidentemente \mathbb{R}^2 , e la funzione è differenziabile in ogni punto del piano. Pertanto, per il Teorema di Fermat i punti di massimo e minimo coincidono con i punti critici, ovvero con le soluzioni del sistema:

$$\begin{cases} f_x = 6x - 3x^2y = 0 \\ f_y = 2y - x^3 = 0 \end{cases}$$

La prima equazione ha per soluzioni $x = 0$ e $xy = 2$: sostituendo nella seconda, si trovano i tre punti critici:

$$(0,0) \quad (\sqrt{2},\sqrt{2}) \quad (-\sqrt{2}, -\sqrt{2}).$$

Resta da discutere la natura di ognuno di questi. L'analisi delle proprietà locali di una funzione di più variabili può effettuarsi, in analogia al caso delle funzioni di variabile reale, sulla base della conoscenza delle derivate parziali seconde. \blacklozenge

Definizione 1.2.3 Data una funzione $f: A \subset \mathbb{R}^n \rightarrow \mathbb{R}$, la sua *matrice hessiana* nel punto $\mathbf{x}_0 \in A$ $H_f(\mathbf{x}_0)$ è una matrice quadrata in cui l'elemento i,j coincide con la derivata seconda di f rispetto alle variabili x_i e x_j :

$$(H_f(\mathbf{x}_0))_{i,j} = f_{x_i x_j}(\mathbf{x}_0).$$

Se le condizioni del Teorema di Schwarz 1.1.7 sono soddisfatte, la matrice hessiana è simmetrica, e quindi tutti i suoi autovalori sono reali.

Esempio 1.2.2

Determinare la matrice hessiana della funzione dell'Esempio 1.2.1. Si ha:

$$\frac{\partial^2 f}{\partial x \partial x} = 6 - 6xy, \quad \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = -3x^2, \quad \frac{\partial^2 f}{\partial y \partial y} = 2,$$

pertanto:

$$H_f = \begin{bmatrix} 6 - 6xy & -3x^2 \\ -3x^2 & 2 \end{bmatrix}.$$

\blacklozenge

Definizione 1.2.4 Data una matrice A di dimensione n , essa è:

1. *definita positiva* (risp. *semidefinita positiva*) se tutti i suoi autovalori sono strettamente positivi (risp. non negativi);
2. *definita negativa* (risp. *semidefinita negativa*) se tutti i suoi autovalori sono strettamente negativi (risp. non positivi);
3. *indefinita* se ha almeno un autovalore strettamente positivo ed uno strettamente negativo.

La determinazione della natura dei punti critici di una funzione di più variabili può essere ricondotta ad una analisi dei segni degli autovalori della matrice hessiana della funzione, valutata nel punto stazionario stesso.

Prima di enunciare il corrispondente teorema, si noti come un punto critico possa non essere nè un punto di massimo nè un punto di minimo (così come per le funzioni di una variabile un punto in cui si annulla la derivata può essere un minimo, un massimo o un flesso):

Definizione 1.2.5 Un punto critico che non sia nè un massimo nè un minimo viene detto *punto di sella*.

Da un punto di vista geometrico, un punto di sella \mathbf{x}_0 per una funzione $f(\mathbf{x})$ è un punto in cui due restrizioni di f su due curve non tangenti in \mathbf{x}_0 presentano un minimo e un massimo (Fig. 1.1).

Teorema 1.2.2 Sia $f: A \subset \mathbb{R}^n \rightarrow \mathbb{R}$, e sia $\mathbf{x}_0 \in A$ un punto critico per f nel quale la funzione sia derivabile.

1. Se $H_f(\mathbf{x}_0)$ è definita positiva, allora \mathbf{x}_0 è un punto di minimo relativo;
2. se $H_f(\mathbf{x}_0)$ è definita negativa, allora \mathbf{x}_0 è un punto di massimo relativo;
3. se $H_f(\mathbf{x}_0)$ è indefinita, allora \mathbf{x}_0 è un punto di sella;
4. se non è verificata nessuna delle condizioni precedenti, allora la natura del punto critico non è determinabile per questa via.

Nel caso particolare di una funzione di due variabili, si ha:

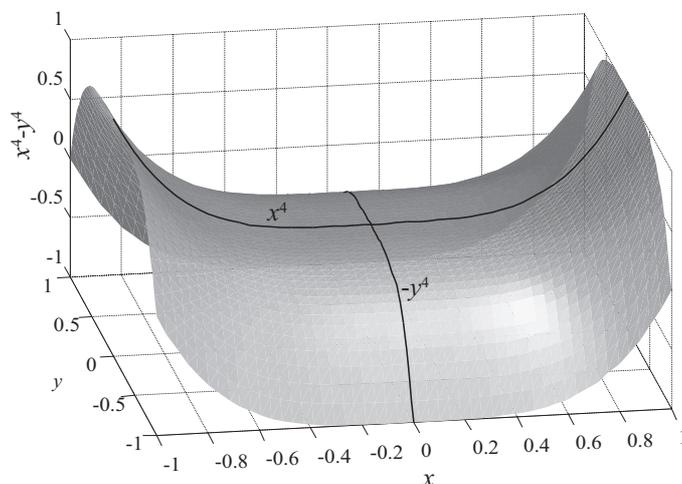


Figura 1.1. Interpretazione geometrica di un punto di sella per la funzione $f(x,y) = x^4 - y^4$.

Teorema 1.2.3 Sia $f: A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$, e sia $\mathbf{x}_0 \in A$ un punto critico per f nel quale la funzione sia derivabile.

1. Se $\det \{H_f(\mathbf{x}_0)\} > 0$ e $f_{x_1 x_1}(\mathbf{x}_0) > 0$, allora \mathbf{x}_0 è un punto di minimo relativo;
2. se $\det \{H_f(\mathbf{x}_0)\} > 0$ e $f_{x_1 x_1}(\mathbf{x}_0) < 0$, allora \mathbf{x}_0 è un punto di massimo relativo;
3. se $\det \{H_f(\mathbf{x}_0)\} < 0$, allora \mathbf{x}_0 è un punto di sella.
4. se $\det \{H_f(\mathbf{x}_0)\} = 0$, allora la natura del punto critico non è determinabile per questa via;

Esempio 1.2.3

Possiamo, a questo punto, completare l'analisi della natura dei punti critici della funzione nell'Esempio 1.2.1. Sulla base della matrice hessiana valutata nell'Esercizio 1.2.2, nei tre punti critici si ha:

$$H_f(0,0) = \begin{bmatrix} 6 & 0 \\ 0 & 2 \end{bmatrix} \quad H_f(\sqrt{2},\sqrt{2}) = \begin{bmatrix} -6 & -6 \\ -6 & 2 \end{bmatrix} \quad H_f(-\sqrt{2},-\sqrt{2}) = \begin{bmatrix} -6 & -6 \\ -6 & 2 \end{bmatrix},$$

pertanto il punto $(0,0)$ è un punto di minimo, mentre $(\sqrt{2},\sqrt{2})$ e $(-\sqrt{2},-\sqrt{2})$ sono due punti di sella. Infatti:

$$\begin{vmatrix} 6 & 0 \\ 0 & 2 \end{vmatrix} = 12 > 0, \quad f_{xx}(0,0) = 6 > 0$$

e

$$\begin{vmatrix} -6 & -6 \\ -6 & 2 \end{vmatrix} = -48 < 0.$$

◆

Esempio 1.2.4

Studiare i punti critici di:

$$f(x,y) = -x^3 + 3xy - 9y^2.$$

Il gradiente della funzione ha componenti:

$$f_x = -3x^2 + 3y \quad f_y = 3x - 18y$$

per cui i punti critici sono le soluzioni del sistema:

$$\begin{cases} -3x^2 + 3y = 0 \\ 3x - 18y = 0 \end{cases} \iff \begin{cases} y = x^2 \\ y = x/6 \end{cases}.$$

Occorre risolvere l'equazione algebrica $x^2 - x/6 = 0$. Si hanno, quindi, i due punti critici:

$$\mathbf{x}_0 = (0,0) \quad \mathbf{x}_1 = \left(\frac{1}{6}, \frac{1}{36}\right).$$

Le componenti della matrice hessiana sono le derivate parziali seconde:

$$f_{xx} = -6x \quad f_{xy} = 3 \quad f_{yy} = -18,$$

per cui:

$$\mathbf{H}_f(\mathbf{x}_0) = \begin{bmatrix} 0 & 3 \\ 3 & -18 \end{bmatrix} \quad \mathbf{H}_f(\mathbf{x}_1) = \begin{bmatrix} -1 & 3 \\ 3 & -18 \end{bmatrix}.$$

I due determinanti valgono:

$$\det \{\mathbf{H}_f(\mathbf{x}_0)\} = -9 < 0 \quad \det \{\mathbf{H}_f(\mathbf{x}_1)\} = 18 - 9 = 9 > 0,$$

e quindi \mathbf{x}_0 è un punto di sella, mentre \mathbf{x}_1 è un punto di massimo, infatti $f_{xx}(\mathbf{x}_1) = -1 < 0$. ◆

Esempio 1.2.5

Studiare i punti critici di:

$$f(x,y) = x^2 + y^2 + x^3y.$$

Il gradiente della funzione ha componenti:

$$f_x = 2x + 3x^2y \quad f_y = 2y + x^3$$

per cui i punti critici sono le soluzioni del sistema:

$$\begin{cases} 2x + 3x^2y = 0 \\ 2y + x^3 = 0 \end{cases}.$$

Dalla seconda, si ricava $y = -x^3/2$, che sostituita nella prima conduce all'equazione algebrica $2x - 3x^5/2 = 0$. Si hanno, quindi, le soluzioni:

$$\mathbf{x}_0 = (0,0) \quad \mathbf{x}_1 = \left(\sqrt[4]{\frac{4}{3}}, -\frac{1}{2} \left(\frac{4}{3} \right)^{3/4} \right) \quad \mathbf{x}_2 = \left(-\sqrt[4]{\frac{4}{3}}, \frac{1}{2} \left(\frac{4}{3} \right)^{3/4} \right).$$

Le componenti della matrice hessiana sono le derivate parziali seconde:

$$f_{xx} = 2 + 6xy \quad f_{xy} = 3x^2 \quad f_{yy} = 2,$$

per cui:

$$\mathbf{H}_f(\mathbf{x}_0) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \quad \mathbf{H}_f(\mathbf{x}_1) = \mathbf{H}_f(\mathbf{x}_2) = \begin{bmatrix} -2 & 3\sqrt{4/3} \\ 3\sqrt{4/3} & 2 \end{bmatrix}.$$

I tre determinanti valgono:

$$\det \{\mathbf{H}_f(\mathbf{x}_0)\} = 4 > 0 \quad \det \{\mathbf{H}_f(\mathbf{x}_1)\} = \det \{\mathbf{H}_f(\mathbf{x}_2)\} = -4 - 12 = -16 < 0,$$

e quindi \mathbf{x}_1 e \mathbf{x}_2 sono punti di sella, mentre \mathbf{x}_0 è un punto di minimo, infatti $f_{xx}(\mathbf{x}_1) = 2 > 0$. ◆

Esempio 1.2.6

Studiare i punti critici della funzione di tre variabili:

$$f(x,y,z) = 3x^2 + 2y^2 + z^2 - 2xz + 2x + 2y + 1.$$

Il gradiente della funzione ha componenti:

$$f_x = 6x + 2 - 2z \quad f_y = 4y + 2 \quad f_z = 2z - 2x$$

per cui i punti critici sono le soluzioni del sistema lineare:

$$\begin{cases} 6x + 2 - 2z = 0 \\ 4y + 2 = 0 \\ 2z - 2x = 0 \end{cases}.$$

Si ricava immediatamente che l'unica soluzione è:

$$\mathbf{x}_0 = \left(-\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2} \right).$$

Le componenti della matrice hessiana sono le derivate parziali seconde:

$$f_{xx} = 6 \quad f_{xy} = 0 \quad f_{xz} = -2 \quad f_{yy} = 4 \quad f_{zz} = 2 \quad f_{yz} = 0,$$

per cui:

$$\mathbf{H}_f(\mathbf{x}_0) = \begin{bmatrix} 6 & 0 & -2 \\ 0 & 4 & 0 \\ -2 & 0 & 2 \end{bmatrix}.$$

Per determinare la natura del punto critico occorre valutare gli autovalori della matrice hessiana, ovvero le soluzioni dell'equazione:

$$\det \{ \mathbf{H}_f(\mathbf{x}_0) - \lambda \mathbf{I} \} = (4 - \lambda) [(6 - \lambda)(2 - \lambda) - 4] = (4 - \lambda)(\lambda^2 - 8\lambda + 8) = 0.$$

Si ha:

$$\lambda_1 = 4 > 0 \quad \lambda_{2,3} = 4 \pm 2\sqrt{2} > 0,$$

e quindi la matrice hessiana è definita positiva. Per il Teorema 1.2.2, \mathbf{x}_0 è un punto di minimo. ◆

Esempio 1.2.7

Studiare i punti critici della funzione:

$$f(x, y) = x^3 + y^3 - (1 + x + y)^3.$$

il gradiente di f ha componenti:

$$f_x = 3x^2 - 3(1 + x + y)^2 \quad f_y = 3y^2 - 3(1 + x + y)^2,$$

per cui i punti critici sono soluzioni del sistema:

$$\begin{cases} 3x^2 - 3(1 + x + y)^2 = 0 \\ 3y^2 - 3(1 + x + y)^2 = 0 \end{cases},$$

ovvero i punti del piano che soddisfano la condizione $x^2 = y^2$. Pertanto, i punti stazionari di f sono tutti i punti di \mathbb{R}^2 che si trovano sulle due rette bisettrici $y = \pm x$.

Le componenti della matrice hessiana sono le derivate seconde:

$$f_{xx} = 6x - 6(1 + x + y) \quad f_{xy} = -6(1 + x + y) \quad f_{yy} = 6y - 6(1 + x + y)$$

per cui:

$$\mathbf{H}_f(x, x) = \begin{bmatrix} -6(1+x) & -6(1+2x) \\ -6(1+2x) & -6(1+x) \end{bmatrix} \quad \mathbf{H}_f(x, -x) = \begin{bmatrix} -6(1-x) & -6 \\ -6 & -6(1+x) \end{bmatrix}.$$

Studiamo separatamente i due casi:

1. punti critici $y = x$. In questo caso, si ha:

$$\det \{\mathbf{H}_f(x, x)\} = -36x(3x + 2),$$

che si annulla in $x = 0$ e $x = -2/3$ (si veda la Fig. 1.2). Pertanto, in questi due punti non possiamo dire nulla sulla natura dei punti stazionari. Invece, per $x < -2/3$ e $x > 0$ il determinante è negativo, e quindi i punti $y = x$ sono punti di sella. Infine, per $-2/3 < x < 0$ il determinante è positivo e $f_{xx}(x, x) = -6(1+x) < 0$, pertanto si tratta di punti di massimo.

2. punti critici $y = -x$. In questo caso, si ha:

$$\det \{\mathbf{H}_f(x, -x)\} = -36x^2,$$

che si annulla solo per $x = 0$, punto stazionario sul quale non possiamo dire nulla. Poiché il determinante è negativo per tutti gli altri valori di x , possiamo concludere che la retta $y = -x$ costituisce, a meno del punto $x = 0$, un insieme di punti di sella.



Esempio 1.2.8

Determinare massimi e minimi per la funzione:

$$f(x, y) = \sqrt{-x^2 - y^2 - xy + 3}.$$

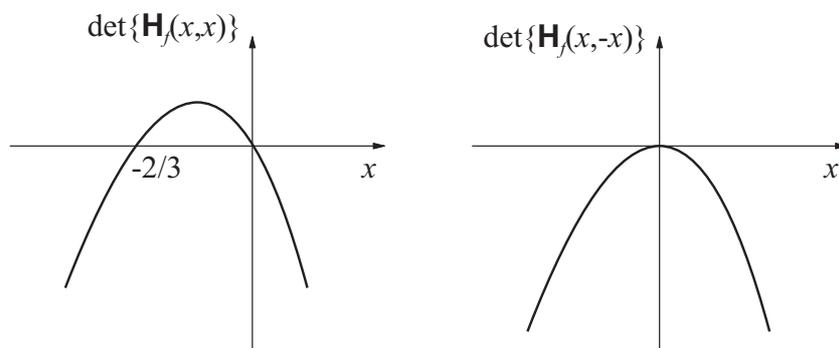


Figura 1.2. Rappresentazione grafica di $\det\{\mathbf{H}_f(x,x)\}$ (a sinistra) e $\det\{\mathbf{H}_f(x,-x)\}$ (a destra) per la funzione dell'Esempio 1.2.7.

Poiché la funzione \sqrt{t} è monotona crescente in t , basta cercare massimi e minimi dell'argomento della radice $g(x,y) = -x^2 - y^2 - xy + 3$ per poter concludere che essi sono punti di massimo e minimo anche per la funzione f . Si ha:

$$g_x = -2x - y \quad g_y = -2y - x,$$

per cui l'unico punto critico è $(0,0)$. La matrice hessiana ha componenti:

$$g_{xx} = -2 \quad g_{xy} = -1 \quad g_{yy} = -2$$

per cui $\det\{\mathbf{H}_g(0,0)\} = 3 > 0$ e $g_{xx}(0,0) = -2 < 0$. Allora $(0,0)$ è un punto di massimo di g , e quindi anche di f . \blacklozenge

1.3 Estremi vincolati

Nella Sez. 1.2 abbiamo studiato le caratteristiche dei punti di massimo e minimo di una funzione di più variabili in tutto il dominio della funzione stessa. Di grande importanza applicativa è il caso di ricerca di estremi di una funzione di più variabili *vincolata* ad un opportuno sottoinsieme del dominio di definizione. In altri termini, la ricerca di estremi vincolati corrisponde alla ricerca dei punti di minimo e massimo di una opportuna restrizione della funzione originale. Per semplicità, nel seguito ci limiteremo a considerare funzioni definite nel piano.

Naturalmente, il caso più semplice è quello in cui il vincolo sia una curva $\Phi : I \subset \mathbb{R} \rightarrow \mathbb{R}^2$ espressa in forma parametrica (ovvero una funzione vettoriale esplicita $\Phi(t) = (x(t), y(t))$), o addirittura espressa da una relazione esplicita tra le variabili

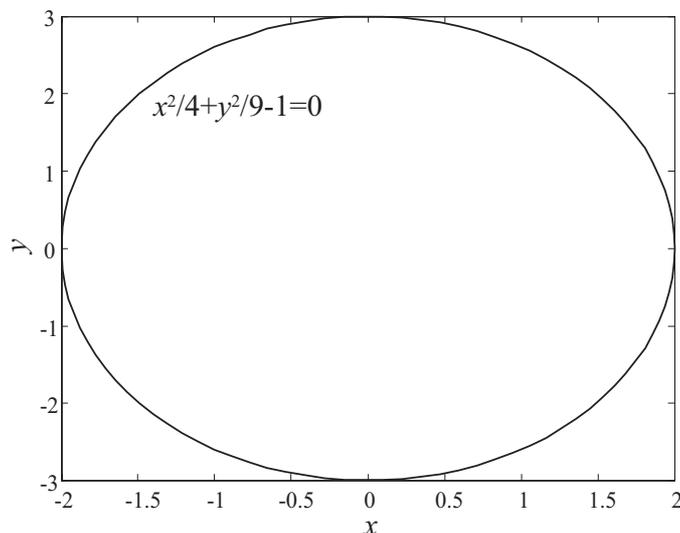


Figura 1.3. Rappresentazione grafica del vincolo $g(x,y) = x^2/4 + y^2/9 - 1 = 0$.

x e y del piano, ad esempio $y = g(x)$. In questo caso, la restrizione della funzione $f(x,y)$ può essere resa esplicita ($f_1(t) = f(x(t),y(t))$ oppure $f_2(x) = f(x,g(x))$) e quindi il problema della ricerca degli estremi vincolati si traduce in una ricerca di massimi e minimi di una funzione di una sola variabile.

Esempio 1.3.1

Si studino gli estremi della funzione $f(x,y) = x^2 + 3y$ con il vincolo $g(x,y) = x^2/4 + y^2/9 - 1 = 0$.

Il vincolo rappresenta un'ellisse nel piano \mathbb{R}^2 di centro l'origine (si veda la Fig. 1.3). Questo può essere facilmente rappresentato in forma parametrica esprimendo:

$$g(x,y) = \left(\frac{x}{2}\right)^2 + \left(\frac{y}{3}\right)^2 - 1 = 0$$

che è soddisfatta se:

$$\frac{x}{2} = \cos t \quad \frac{y}{3} = \sin t \quad 0 \leq t \leq 2\pi,$$

ovvero definendo la curva $\Phi(t) = (2 \cos t, 3 \sin t)$ per $t \in I = \{x \in \mathbb{R} : 0 \leq x \leq 2\pi\}$. La restrizione della funzione f al vincolo è quindi:

$$f_1(t) = f(\Phi(t)) = 4 \cos^2 t + 9 \sin t,$$

i cui estremi soddisfano:

$$\frac{df_1}{dt} = \cos t (-8 \sin t + 9) = 0.$$

Il secondo fattore non si annulla mai per $t \in \mathbb{R}$ (infatti ciò richiederebbe $\sin t = 9/8 > 1$), per cui gli estremi sono le soluzioni di:

$$\cos t = 0 \text{ per } t \in I \iff t = \frac{\pi}{2}, \frac{3\pi}{2}.$$

Nel piano (x, y) si sono quindi individuati i due punti $\mathbf{x}_1 = (0, 3)$ e $\mathbf{x}_2 = (0, -3)$. Per classificare la natura dei due punti critici, occorre studiare il segno della derivata seconda:

$$\frac{d^2 f_1}{dt^2} = 8 \sin^2 t - 8 \cos^2 t - 9 \sin t,$$

che vale -1 in \mathbf{x}_1 e 17 in \mathbf{x}_2 . Pertanto, \mathbf{x}_1 è un punto di massimo per la restrizione considerata, e \mathbf{x}_2 un punto di minimo. \blacklozenge

In pratica, il vincolo raramente può essere espresso in forma esplicita: nel caso generale, si può sfruttare il seguente

Teorema 1.3.1 Sia $\mathbf{x}_0 = (x_0, y_0)$ un punto di estremo vincolato per $f(x, y)$ rispetto al vincolo $g(x, y) = 0$. Se $\nabla g(\mathbf{x}_0) \neq \mathbf{0}$, allora esiste un numero $\lambda_0 \in \mathbb{R}$:

$$\nabla f(\mathbf{x}_0) = \lambda_0 \nabla g(\mathbf{x}_0).$$

Pertanto, si possono ricercare gli estremi di f vincolati a g seguendo i passi (*metodo dei moltiplicatori di Lagrange*):

1. si costruisce la funzione di tre variabili (detta *lagrangiana*):

$$L(x, y, \lambda) = f(x, y) - \lambda g(x, y);$$

2. si determinano i punti critici liberi (x_0, y_0, λ_0) della lagrangiana, che per definizione soddisfano il sistema:

$$\begin{cases} L_x = f_x(x, y) - \lambda g_x(x, y) = 0 \\ L_y = f_y(x, y) - \lambda g_y(x, y) = 0 \\ L_\lambda = -g(x, y) = 0 \end{cases},$$

e quindi, purché $\nabla g(x_0, y_0) \neq \mathbf{0}$, soddisfano sia il Teorema 1.3.1 (prime due equazioni del sistema) sia la condizione di vincolo (ultima equazione del sistema);

3. si determina la natura dei punti critici trovati, spesso facendo uso del Teorema di Weierstrass 1.1.6.

Esempio 1.3.2

Risolvere il quesito dell'Esempio 1.3.1 facendo uso del metodo dei moltiplicatori di Lagrange.

Le due funzioni $f(x,y)$ e $g(x,y)$ sono entrambe di classe $C^\infty(\mathbb{R}^2)$, e la lagrangiana si scrive:

$$L(x,y,\lambda) = x^2 + 3y - \lambda \left(\frac{x^2}{4} + \frac{y^2}{9} - 1 \right),$$

i cui punti critici soddisfano il sistema:

$$\begin{cases} L_x = 2x - \lambda x/2 = 0 \\ L_y = 3 - \lambda 2y/9 = 0 \\ L_\lambda = -(x^2/4 + y^2/9 - 1) = 0 \end{cases}.$$

La prima equazione ha le due soluzioni:

$$x = 0 \quad \lambda = 4,$$

delle quali, la seconda conduce a una soluzione non reale del sistema complessivo. Infatti, per $\lambda = 4$, dalla seconda equazione del sistema è:

$$y = \frac{27}{2\lambda} = \frac{27}{8},$$

mentre la terza equazione si scrive:

$$x^2 = 4(1 - y^2/9) = 4(1 - 81/64) < 0$$

che quindi non ha soluzioni reali.

Resta da analizzare il solo caso $x = 0$: dalla terza equazione si ricava $y = \pm 3$, e quindi dalla seconda: $\lambda = \pm 9/2$. In definitiva, i punti critici della lagrangiana sono $(\mathbf{x}_1, \lambda_1) = (0, 3, 9/2)$ e $(\mathbf{x}_2, \lambda_2) = (0, -3, -9/2)$. Il vincolo definisce un insieme chiuso, limitato e non vuoto in \mathbb{R}^2 , mentre il Teorema 1.3.1 garantisce che \mathbf{x}_1 e \mathbf{x}_2 sono estremi di f vincolata a g (infatti $\nabla g(\mathbf{x}_1) = (0, 2/3) \neq \mathbf{0}$ e $\nabla g(\mathbf{x}_2) = (0, -2/3) \neq \mathbf{0}$). Pertanto, essendo $f(\mathbf{x}_1) = 9$ e $f(\mathbf{x}_2) = -9$, per il Teorema di Weierstrass \mathbf{x}_1 e \mathbf{x}_2 sono, rispettivamente, un punto di massimo e di minimo per f vincolata a g . \blacklozenge

2

Equazioni differenziali

Obiettivi: Equazioni differenziali scalari del primo ordine a variabili separabili. Equazioni differenziali scalari lineari del primo ordine. Sistemi di equazioni differenziali lineari del primo ordine a coefficienti costanti. Equazioni differenziali scalari lineari a coefficienti costanti di ordine superiore a uno.

2.1 Introduzione

Sia y una funzione della variabile reale t , e sia F una funzione reale delle $n + 2$ variabili $t, y, y', \dots, y^{(n)}$, dove:

$$y^{(n)} = \frac{d^n y}{dt^n}.$$

Definizione 2.1.1 Si chiama *equazione differenziale scalare* nella variabile y la relazione:

$$F(t, y(t), y'(t), \dots, y^{(n)}(t)) = 0. \quad (2.1)$$

L'*ordine* della equazione differenziale coincide con il massimo ordine di derivazione presente nella funzione F .

Una funzione $\varphi(t): I \subset \mathbb{R} \rightarrow \mathbb{R}$, definita e derivabile n volte in un intervallo I , che soddisfi la (2.1) $\forall t \in I$ viene detta *soluzione* dell'equazione differenziale. Si può dimostrare che una equazione differenziale di ordine n è soddisfatta da un insieme infinito di funzioni di t caratterizzato da n *parametri* $p_1, \dots, p_n \in \mathbb{R}$, ovvero la (2.1) è risolta da una *famiglia* di funzioni $\varphi(t; p_1, \dots, p_n)$; tale famiglia viene detta *soluzione generale* dell'equazione differenziale in esame. Per selezionare, come caso particolare

della soluzione generale, una *soluzione particolare* occorre assegnare n condizioni che la soluzione particolare stessa deve soddisfare, in modo da definire univocamente i parametri $p_{1,0}, \dots, p_{n,0}$. Ad esempio, si possono assegnare le n condizioni iniziali:

$$y(t_0) = y_0 \quad y'(t_0) = y'_0 \quad \dots \quad y^{(n-1)}(t_0) = y_0^{(n-1)} \quad t_0 \in I. \quad (2.2)$$

La soluzione particolare cercata sarà quindi $\varphi(t; p_{1,0}, \dots, p_{n,0})$.

Una equazione differenziale di ordine n viene detta in *forma normale* se può essere espressa secondo:

$$y^{(n)}(t) = f(t, y(t), y'(t), \dots, y^{(n-1)}(t)). \quad (2.3)$$

Definizione 2.1.2 Una equazione differenziale in forma normale completata dalle condizioni iniziali (2.2) viene detta *problema di Cauchy*.

Qualora la funzione f soddisfi ben precise condizioni di regolarità, è possibile dimostrare che il problema di Cauchy ammette soluzione unica, almeno in un intorno di t_0 .

2.2 Equazioni differenziali a variabili separabili

Si tratta di equazioni differenziali scalari del primo ordine che possono essere espresse nella forma:

$$y'(t) = a(t)b(y(t)), \quad (2.4)$$

dove $a: I \subset \mathbb{R} \rightarrow \mathbb{R}$ e $b: J \subset \mathbb{R} \rightarrow \mathbb{R}$ sono funzioni reali continue. Si noti che la soluzione \bar{y} dell'equazione algebrica:

$$b(\bar{y}) = 0$$

è soluzione dell'equazione a variabili separabili (2.4), poiché $d\bar{y}/dt = 0$. Altre soluzioni possono essere trovate in $J' \subseteq J$ dove $b(y) \neq 0$,¹ scrivendo (2.4) nella forma:

$$\frac{y'}{b(y)} = a(t),$$

dalla quale si deduce immediatamente:

$$\int \frac{y'(t)}{b(y(t))} dt = \int a(t) dt + c,$$

¹ovvero, $J' = \{y \in J: b(y) \neq 0\}$.

dove $c \in \mathbb{R}$ è una costante arbitraria e

$$\int g(t) dt$$

indica una qualsiasi primitiva della funzione $g(t)$, ovvero una funzione $G(t)$ tale che $G'(t) = g(t)$. Se a primo membro si fa uso del cambiamento di variabile $y = y(t) \Rightarrow dy = y'(t)dt$ si ottiene:

$$\int \frac{1}{b(y)} dy = \int a(t) dt + c. \quad (2.5)$$

Se $B(y)$ è una primitiva di $1/b(y)$, e $A(t)$ una primitiva di $a(t)$, la (2.5) definisce implicitamente la famiglia di soluzioni della (2.4):

$$B(y) = A(t) + c.$$

Teorema 2.2.1 Si consideri il problema di Cauchy:

$$\begin{cases} y' = a(t)b(y) \\ y(t_0) = y_0 \end{cases}$$

definito nell'intervallo I e con $t_0 \in I$. Se la funzione $a(t)$ è continua in un intorno di t_0 e $b(y)$ è di classe C^1 in un intorno di y_0 , allora il problema di Cauchy ha soluzione unica, almeno in un intorno di t_0 .

Esempio 2.2.1

Risolvere il problema di Cauchy:

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

Si tratta di una equazione differenziale a variabili separabili, nella forma (2.4) dove $b(y) = 1/y$ e $a(x) = 1$. Visto che le due funzioni soddisfano le condizioni del Teorema 2.2.1 in un intorno della condizione iniziale $x_0 = 0$ e $y_0 = 2$, il problema di Cauchy ammette soluzione unica. Poiché $b(y) \neq 0 \forall y$, non vi sono soluzioni particolari del tipo $y = \text{cost}$. Allora, la soluzione generale è data dalla (2.5), dove:

$$\begin{aligned} \int \frac{1}{b(y)} dy &= \int y dy = \frac{1}{2}y^2 \\ \int a(x) dx &= \int dx = x, \end{aligned}$$

cioè:

$$\frac{1}{2}y^2 = x + c \implies y = \pm\sqrt{2(x+c)},$$

definita per $x \geq -c$. Imponendo la condizione iniziale:

$$y_0 = \pm\sqrt{2(x_0+c)} \implies 2 = \pm\sqrt{2c}$$

si deduce immediatamente che la soluzione del problema di Cauchy corrisponde alla determinazione positiva della radice e a $c = 2$, ovvero:

$$y(x) = \sqrt{2x+4},$$

definita per $x \geq -2$. ◆

Esempio 2.2.2

Determinare le soluzioni dell'equazione differenziale:

$$(t^2 + 1)y' + y^2 = 0.$$

Si tratta di una equazione differenziale a variabili separabili, infatti può essere espressa nella forma:

$$y' = -\frac{y^2}{1+t^2},$$

che è equivalente alla forma originale poiché $1+t^2 \neq 0 \forall t \in \mathbb{R}$. Una soluzione particolare è:

$$b(y) = y^2 = 0 \implies y = \bar{y} = 0,$$

mentre altre soluzioni sono date da:

$$\int \frac{dy}{y^2} = -\int \frac{dt}{1+t^2} + c,$$

dove:

$$\int \frac{dy}{y^2} = -\frac{1}{y}, \quad -\int \frac{dt}{1+t^2} = -\arctan t.$$

Pertanto:

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

rappresenta una famiglia di soluzioni dell'equazione differenziale in esame, cui si aggiunge la soluzione particolare $y = 0$. ◆

Esempio 2.2.3

Risolvere il problema di Cauchy:

$$\begin{cases} y' + x \tan y = 0 \\ y(0) = \frac{\pi}{2} \end{cases} .$$

Separando le variabili, si ottiene:

$$\frac{y'}{\tan y} = -x$$

per cui, essendo:

$$\int \frac{dy}{\tan y} = \int \frac{\cos y}{\sin y} dy = \ln(\sin y), \quad - \int x dx = -\frac{1}{2}x^2$$

una famiglia di soluzioni dell'equazione differenziale è:

$$\ln(\sin y) = -\frac{1}{2}x^2 + c \implies \sin y = ae^{-x^2/2} \quad a = e^c \in \mathbb{R}^+,$$

completata dalle soluzioni particolari: $\tan y = 0 \implies y = k\pi \quad k \in \mathbb{Z}$. Imponendo la condizione iniziale:

$$\sin y(0) = \sin \frac{\pi}{2} = 1 = a,$$

si ricava la soluzione del problema di Cauchy:

$$\sin y = e^{-x^2/2} \implies y = \arcsin \left(e^{-x^2/2} \right),$$

che risulta essere definita per $x \in \mathbb{R}$ poiché in questo dominio $0 < e^{-x^2/2} < 1$. \blacklozenge

2.3 Equazioni differenziali lineari del primo ordine

Un'equazione differenziale in forma normale (2.3) è *lineare* se la funzione f è lineare nelle variabili $y, y', y^{(n-1)}$, ovvero se, date due soluzioni particolari $y_1(t)$ e $y_2(t)$, si ha:

$$\begin{aligned} f(t, \alpha y_1 + \beta y_2, \alpha y_1' + \beta y_2', \dots, \alpha y_1^{(n-1)} + \beta y_2^{(n-1)}) \\ = \alpha f(t, y_1, y_1', \dots, y_1^{(n-1)}) + \beta f(t, y_2, y_2', \dots, y_2^{(n-1)}) \quad \alpha, \beta \in \mathbb{R}. \end{aligned}$$

In altri termini, qualunque combinazione lineare di due soluzioni particolari è ancora soluzione della stessa equazione differenziale. Nella sua forma più generale, allora, una equazione differenziale lineare di ordine n si può esprimere nella forma:

$$y^{(n)}(t) + \sum_{k=0}^{n-1} a_k(t)y^{(k)}(t) = s(t), \quad (2.6)$$

dove $a_k(t)$ e $s(t)$ sono funzioni note continue su $I \subset \mathbb{R}$. $s(t)$ viene spesso detto *termine di sorgente*, *termine noto* o *termine forzante* dell'equazione differenziale. Nel caso in cui $s(t) = 0$, si parla di *equazione omogenea*, mentre nel caso opposto l'equazione è *completa*.

Definizione 2.3.1 Data una equazione differenziale lineare del tipo (2.6), la sua *omogenea associata* è l'equazione differenziale che si ottiene ponendo $s(t) = 0$.

Teorema 2.3.1 L'integrale generale di una equazione differenziale lineare si ottiene sommando all'integrale generale della equazione omogenea associata una soluzione particolare dell'equazione completa.

In particolare, ci occuperemo in questo contesto di equazioni differenziali lineari del primo ordine, ovvero di equazioni della forma:

$$y'(t) + a(t)y(t) = s(t). \quad (2.7)$$

Per questa classe di equazioni differenziali, è possibile costruire una forma esplicita dell'integrale generale facendo uso del Teorema 2.3.1:

Omogenea associata. Si tratta di risolvere l'equazione differenziale:

$$z'(t) + a(t)z(t) = 0.$$

Se $A(t)$ è una primitiva di $a(t)$, ovvero una funzione tale che $A'(t) = a(t)$, moltiplicando l'equazione omogenea per $e^{A(t)}$ si ottiene:

$$z'(t)e^{A(t)} + a(t)z(t)e^{A(t)} = [z(t)e^{A(t)}]' = 0,$$

da cui è immediato ricavare:

$$z(t)e^{A(t)} = c \quad c \in \mathbb{R},$$

o anche:

$$z(t) = ce^{-A(t)} = ce^{-\int a(t) dt}.$$

Soluzione particolare. Resta da costruire una soluzione particolare dell'equazione completa. Si può procedere seguendo il *metodo di variazione delle costanti*, che consiste nel cercare una soluzione $\bar{y}(t)$ della stessa forma della soluzione dell'equazione omogenea, ma con la costante $c \in \mathbb{R}$ ora supposta funzione (incognita) della variabile di integrazione:

$$\bar{y}(t) = c(t)e^{-A(t)}.$$

Sostituendo nell'equazione completa, si ha:

$$\bar{y}'(t) + a(t)\bar{y}(t) = c'(t)e^{-A(t)} - c(t)a(t)e^{-A(t)} + a(t)c(t)e^{-A(t)} = c'(t)e^{-A(t)} = s(t),$$

da cui:

$$c'(t) = e^{A(t)}s(t) \implies c(t) = \int e^{A(t)}s(t) dt.$$

Pertanto:

$$\bar{y}(t) = e^{-A(t)} \int e^{A(t)}s(t) dt.$$

In definitiva, facendo uso del Teorema 2.3.1, la soluzione generale della (2.7) è data da:

$$y(t) = z(t) + \bar{y}(t) = ce^{-A(t)} + e^{-A(t)} \int e^{A(t)}s(t) dt \quad (2.8a)$$

$$A(t) = \int a(t) dt \quad c \in \mathbb{R}. \quad (2.8b)$$

Se si è in presenza di un problema di Cauchy, caratterizzato dalla (2.7) e da

$$y(t_0) = y_0 \quad t_0 \in I,$$

la soluzione generale (2.8) deve anche soddisfare la condizione iniziale. Scegliendo $A(t) = A_0(t)$, dove $A_0(t)$ è la primitiva di $a(t)$ che si annulla in t_0 :

$$A_0(t) = \int_{t_0}^t a(t) dt,$$

la (2.8) soddisfa la condizione iniziale se $c = y_0$ e se anche la primitiva di $s(t) \exp A(t)$ si annulla per $t = t_0$, e quindi la soluzione del problema di Cauchy si scrive:

$$y(t) = y_0 e^{-A_0(t)} + e^{-A_0(t)} \int_{t_0}^t e^{A_0(t)} s(t) dt. \quad (2.9)$$

Esempio 2.3.1

Risolvere il problema di Cauchy:

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

Si tratta di una equazione differenziale lineare del primo ordine del tipo (2.7), dove:

$$a(x) = \frac{2}{x} \quad s(x) = \frac{1}{x^2}$$

sono continue in $\mathbb{R} \setminus \{0\}$. Visto che il problema di Cauchy ha condizione iniziale in $x_0 = -1$, si considera l'intervallo $I =]-\infty, 0[$. La primitiva di $a(x)$ che si annulla in x_0 è:

$$A_0(x) = \int_{-1}^x a(x) \, dx = \int_{-1}^x \frac{2}{x} \, dx = [2 \ln |x|]_{-1}^x = 2 \ln(-x) \quad x \in I.$$

Resta da valutare l'integrale a secondo membro della (2.9):

$$\int_{-1}^x e^{A_0(x)} s(x) \, dx = \int_{-1}^x \frac{1}{x^2} e^{2 \ln(-x)} \, dx = \int_{-1}^x dx = x + 1.$$

La soluzione del problema di Cauchy è quindi data dalla (2.9):

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

◆

Esempio 2.3.2

Risolvere il problema di Cauchy:

$$\begin{cases} x' - x = 1 \\ x(0) = 0 \end{cases} .$$

L'equazione differenziale è lineare con $a(t) = -1$ e $s(t) = 1$, per cui:

$$A_0(t) = \int_0^t a(t) \, dt = - \int_0^t dt = -t$$

$$\int_0^t e^{A_0(t)} s(t) \, dt = \int_0^t e^{-t} \, dt = 1 - e^{-t}.$$

La soluzione è quindi:

$$x(t) = 0 + e^t (1 - e^{-t}) = e^t - 1.$$

◆

Esempio 2.3.3

Trovare la soluzione generale dell'equazione differenziale:

$$(\sin t)x' + (\cos t)x = e^t.$$

Si tratta di una equazione differenziale lineare del primo ordine:

$$x' + \frac{\cos t}{\sin t}x = \frac{e^t}{\sin t},$$

per cui $a(t) = \cos t / \sin t$ e $s(t) = e^t / \sin t$. Allora:

$$A(t) = \int a(t) dt = \int \frac{\cos t}{\sin t} dt = \ln(\sin t) \implies e^{A(t)} = \sin t$$

e inoltre:

$$\int e^{A(t)} s(t) dt = \int \sin t \frac{e^t}{\sin t} dt = \int e^t dt = e^t.$$

La soluzione generale è quindi data dalla (2.8):

$$x(t) = \frac{c}{\sin t} + \frac{e^t}{\sin t} = \frac{c + e^t}{\sin t} \quad c \in \mathbb{R}.$$

**Esempio 2.3.4**

Trovare l'integrale generale dell'equazione differenziale:

$$y' = -2y + e^{-2x}$$

e determinare la soluzione particolare che ha derivata pari a 1 per $x = 0$.

L'equazione lineare del primo ordine è caratterizzata da:

$$a(x) = 2 \quad s(x) = e^{-2x},$$

per cui:

$$A(x) = \int a(x) dx = \int 2 dx = 2x$$

$$\int e^{A(x)} s(x) dx = \int e^{2x} e^{-2x} dx = x.$$

Sostituendo nella (2.8) si ha la soluzione generale:

$$y(x) = ce^{-2x} + e^{-2x}x = e^{-2x}(x + c) \quad c \in \mathbb{R}.$$

La soluzione particolare deve soddisfare la condizione:

$$y'(0) = \left[-2e^{-2x}(c + x) + e^{-2x} \right]_{x=0} = -2c + 1 = 1$$

da cui $c = 0$. La soluzione particolare richiesta è quindi $y(x) = x \exp(-2x)$. ◆

Esempio 2.3.5

Risolvere il problema di Cauchy:

$$\begin{cases} tx' - 2x = t^4 \cos t \\ x(\pi) = 0 \end{cases}.$$

L'equazione differenziale lineare è caratterizzata da $a(t) = -2/t$ e $s(t) = t^3 \cos t$, per cui:

$$\begin{aligned} A_0(t) &= \int_{\pi}^t a(t) dt = - \int_{\pi}^t \frac{2}{t} dt = -2 \ln \frac{t}{\pi} \\ \int_{\pi}^t e^{A_0(t)} s(t) dt &= \int_{\pi}^t \frac{\pi^2}{t^2} t^3 \cos t dt = \pi^2 \int_{\pi}^t t \cos t dt. \end{aligned}$$

Integrando per parti si ha:

$$\int t \cos t dt = t \sin t - \int \sin t dt = t \sin t + \cos t,$$

da cui:

$$\int_{\pi}^t e^{A_0(t)} s(t) dt = \pi^2 (t \sin t + \cos t + 1).$$

Sostituendo nella (2.9) si determina la soluzione:

$$x(t) = \frac{t^2}{\pi^2} \pi^2 (t \sin t + \cos t + 1) = t^3 \sin t + t^2 \cos t + t^2.$$

◆

Si noti che nella definizione precedente A^k indica il prodotto (righe per colonne) di k fattori, ognuno dei quali uguale ad A :

$$A^k = \underbrace{A \cdot A \cdots A}_{k \text{ termini}}$$

Teorema 2.4.1 La serie (2.12) converge qualunque sia la matrice $A \in \mathbb{R}^{n \times n}$.

In particolare, interessa introdurre la funzione $F_A: \mathbb{R} \rightarrow \mathbb{R}^{n \times n}$:

$$F_A(t) = e^{tA},$$

che risulta essere definita su tutto \mathbb{R} per il teorema precedente.

Teorema 2.4.2 Data una matrice reale quadrata $A \in \mathbb{R}^{n \times n}$ e definita la funzione $F_A(t) = \exp(tA)$, si ha che:

1. la funzione F_A è derivabile in \mathbb{R} , e si ha:

$$F'_A(t) = A \cdot e^{tA} = e^{tA} \cdot A;$$

2. valgono le seguenti proprietà:

$$\left(e^{tA}\right)^{-1} = e^{-tA} \quad e^{(t+s)A} = e^{tA} \cdot e^{sA} = e^{sA} \cdot e^{tA}.$$

Consideriamo, ora, il sistema lineare omogeneo associato a (2.11), cioè il sistema differenziale del primo ordine a coefficienti costanti:

$$\mathbf{X}'(t) = A \cdot \mathbf{X}(t). \tag{2.13}$$

Teorema 2.4.3 Le soluzioni del sistema omogeneo (2.13) sono tutte e sole le funzioni:

$$\mathbf{X}(t) = e^{tA} \cdot \mathbf{c},$$

dove $\mathbf{c} \in \mathbb{R}^n$ è un vettore di costanti arbitrarie reali.

Le colonne $\mathbf{X}_j(t)$ della matrice $\exp(tA)$ costituiscono n soluzioni indipendenti del sistema (2.13), e si dice che ne sono un *sistema fondamentale di soluzioni*, nel senso che ogni altra soluzione del sistema può essere ottenuta come combinazione lineare di queste: in altri termini, le colonne di $\exp(tA)$ costituiscono una base di uno spazio vettoriale di dimensione n che, a sua volta, comprende tutte le soluzioni del sistema omogeneo (2.13).

Dal Teorema 2.4.3 segue facilmente:

Teorema 2.4.4 Il problema di Cauchy:

$$\begin{cases} \mathbf{X}' = \mathbf{A} \cdot \mathbf{X} \\ \mathbf{X}(t_0) = \mathbf{X}_0 \end{cases}$$

ammette una e una sola soluzione, pari a:

$$\mathbf{X}(t) = e^{(t-t_0)\mathbf{A}} \cdot \mathbf{X}_0.$$

Per quanto riguarda il sistema completo (2.11), vale:

Teorema 2.4.5 Tutte e sole soluzioni del sistema differenziale del primo ordine (2.11), nel quale $\mathbf{S}(t)$ è costituito da funzioni continue su un intervallo $I \subset \mathbb{R}$, sono espresse da:

$$\mathbf{Y}(t) = e^{t\mathbf{A}} \cdot \left[\mathbf{c} + \int_{t_0}^t e^{-t\mathbf{A}} \cdot \mathbf{S}(t) dt \right], \quad (2.14)$$

dove $\mathbf{c} \in \mathbb{R}^n$ e $t_0 \in I$.

Ovvero, in accordo al Teorema 2.3.1 esteso al caso dei sistemi di equazioni differenziali lineari, la soluzione generale del sistema completo (2.11) si ottiene sommando la soluzione generale del sistema omogeneo (2.13) ad una soluzione particolare del sistema completo, espressa dal secondo addendo della (2.14).

In definitiva, per determinare la soluzione di un sistema differenziale del primo ordine a coefficienti costanti è fondamentale riuscire a calcolare la matrice esponenziale, ovvero sommare la serie (2.12). Vedremo tre casi particolari:

1. La matrice \mathbf{A} è diagonale:

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

In questo caso, si ha:

$$e^{t\mathbf{A}} = \begin{bmatrix} e^{\lambda_1 t} & 0 & 0 & \dots & 0 \\ 0 & e^{\lambda_2 t} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & e^{\lambda_n t} \end{bmatrix}. \quad (2.15)$$

2. La matrice A ha un unico autovalore λ di molteplicità n , che naturalmente deve essere reale. In questo caso, A può essere decomposta in $A = \lambda I + N$, dove I è la matrice identica e N è una matrice la cui serie esponenziale contiene al più i primi n termini:

$$e^{tA} = e^{t\lambda I + tN} = e^{t\lambda I} \cdot e^{tN} = e^{\lambda t} \sum_{k=0}^{n-1} \frac{t^k}{k!} N^k = e^{\lambda t} \sum_{k=0}^{n-1} \frac{t^k}{k!} (A - \lambda I)^k. \quad (2.16)$$

3. La matrice A è diagonalizzabile, ovvero esiste una matrice invertibile B tale che:

$$B^{-1} \cdot A \cdot B = D = \begin{bmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda_n \end{bmatrix},$$

essendo λ_i gli autovalori di A . In questo caso, si può dimostrare che:

$$e^{tA} = e^{tB \cdot D \cdot B^{-1}} = B \cdot e^{tD} \cdot B^{-1}, \quad (2.17)$$

dove la matrice esponenziale della matrice diagonale tD può essere valutata sulla base della (2.15).

Esempio 2.4.1

Trovare la soluzione generale del sistema differenziale:

$$\begin{cases} x_1' = x_1 \\ x_2' = x_1 + x_2 \end{cases}.$$

La matrice dei coefficienti è:

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix},$$

i cui autovalori sono soluzione di:

$$\det \{A - \lambda I\} = \begin{vmatrix} 1 - \lambda & 0 \\ 1 & 1 - \lambda \end{vmatrix} = (1 - \lambda)^2 = 0$$

ovvero, si hanno i due autovalori coincidenti $\lambda_{1,2} = 1$. Pertanto, la matrice esponenziale può essere valutata applicando la (2.16), dove:

$$N = A - I = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

Si ha quindi:

$$e^{t\mathbf{A}} = e^t [\mathbf{I} + t\mathbf{N}] = e^t \begin{bmatrix} 1 & 0 \\ t & 1 \end{bmatrix}.$$

Infine, la soluzione generale cercata è data dal Teorema 2.4.3:

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = e^t \begin{bmatrix} 1 & 0 \\ t & 1 \end{bmatrix} \cdot \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} c_1 e^t \\ c_1 t e^t + c_2 e^t \end{bmatrix}.$$



Esempio 2.4.2

Trovare la soluzione generale del sistema differenziale:

$$\begin{cases} x_1' = 2x_2 \\ x_2' = 2x_1 \end{cases}.$$

La matrice dei coefficienti è:

$$\mathbf{A} = \begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix},$$

per cui gli autovalori sono le soluzioni del polinomio caratteristico:

$$\det \{\mathbf{A} - \lambda \mathbf{I}\} = \begin{vmatrix} -\lambda & 2 \\ 2 & -\lambda \end{vmatrix} = \lambda^2 - 4 = 0$$

ovvero $\lambda_{1,2} = \pm 2$. Visto che i due autovalori sono distinti, la matrice dei coefficienti può essere diagonalizzata da una matrice \mathbf{B} le cui colonne siano autovettori di \mathbf{A} . Allora:

1. autovalore $\lambda_1 = 2$, per cui l'autovettore \mathbf{v}_1 deve soddisfare:

$$[\mathbf{A} - 2\mathbf{I}] \cdot \mathbf{v}_1 = \begin{bmatrix} -2 & 2 \\ 2 & -2 \end{bmatrix} \cdot \begin{bmatrix} v_{11} \\ v_{12} \end{bmatrix},$$

cioè l'equazione $v_{11} = v_{12}$. Pertanto, un autovettore associato a 2 è:

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix};$$

2. autovalore $\lambda_2 = -2$, per cui l'autovettore \mathbf{v}_2 deve soddisfare:

$$[\mathbf{A} + 2\mathbf{I}] \cdot \mathbf{v}_2 = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} \cdot \begin{bmatrix} v_{21} \\ v_{22} \end{bmatrix},$$

cioè l'equazione $v_{21} = -v_{22}$. Pertanto, un autovettore associato a -2 è:

$$\mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

La matrice \mathbf{B} è quindi:

$$\mathbf{B} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \implies \mathbf{B}^{-1} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Infine, per la (2.17):

$$e^{t\mathbf{A}} = \mathbf{B} \cdot \begin{bmatrix} e^{2t} & 0 \\ 0 & e^{-2t} \end{bmatrix} \cdot \mathbf{B}^{-1} = \frac{1}{2} \begin{bmatrix} e^{2t} + e^{-2t} & e^{2t} - e^{-2t} \\ e^{2t} - e^{-2t} & e^{2t} + e^{-2t} \end{bmatrix},$$

e quindi la soluzione generale cercata è:

$$\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = e^{t\mathbf{A}} \cdot \mathbf{c} = \frac{1}{2} \begin{bmatrix} (c_1 + c_2)e^{2t} + (c_1 - c_2)e^{-2t} \\ (c_1 + c_2)e^{2t} - (c_1 - c_2)e^{-2t} \end{bmatrix} = \begin{bmatrix} c'_1 e^{2t} + c'_2 e^{-2t} \\ c'_1 e^{2t} - c'_2 e^{-2t} \end{bmatrix},$$

dove si è posto $c'_1 = (c_1 + c_2)/2 \in \mathbb{R}$ e $c'_2 = (c_1 - c_2)/2 \in \mathbb{R}$. ◆

Esempio 2.4.3

Risolvere il problema di Cauchy:

$$\begin{cases} \mathbf{X}' = \mathbf{A} \cdot \mathbf{X} \\ \mathbf{X}(0) = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \end{cases},$$

dove la matrice dei coefficienti è:

$$\mathbf{A} = \begin{bmatrix} 3 & -1 & -1 \\ 5 & -2 & -4 \\ -4 & 3 & 5 \end{bmatrix}.$$

Gli autovalori di A sono le radici del polinomio caratteristico:

$$\det \{A - \lambda I\} = \begin{vmatrix} 3 - \lambda & -1 & -1 \\ 5 & -2 - \lambda & -4 \\ -4 & 3 & 5 - \lambda \end{vmatrix} = -(\lambda^3 - 6\lambda^2 + 12\lambda - 8) = -(\lambda - 2)^3,$$

e quindi si tratta di tre autovalori coincidenti $\lambda_{1,2,3} = 2$. Si può allora fare uso della (2.16), dove:

$$N = A - 2I = \begin{bmatrix} 1 & -1 & -1 \\ 5 & -4 & -4 \\ -4 & 3 & 3 \end{bmatrix}$$

$$N^2 = N \cdot N = \begin{bmatrix} 0 & 0 & 0 \\ 1 & -1 & -1 \\ -1 & 1 & 1 \end{bmatrix}.$$

Si ha allora:

$$e^{tN} = I + tN + \frac{t^2}{2}N^2 = \begin{bmatrix} t+1 & -t & -t \\ 5t + \frac{t^2}{2} & -4t - \frac{t^2}{2} + 1 & -4t - \frac{t^2}{2} \\ -4t - \frac{t^2}{2} & 3t + \frac{t^2}{2} & 3t + \frac{t^2}{2} + 1 \end{bmatrix},$$

per cui grazie alla (2.16) e al Teorema 2.4.4:

$$\mathbf{X}(t) = e^{tA} \cdot \mathbf{X}_0 = e^{2t} e^{tN} \cdot \mathbf{X}_0 = e^{2t} \begin{bmatrix} t+1 \\ 5t + \frac{t^2}{2} \\ -4t - \frac{t^2}{2} \end{bmatrix}.$$

◆

2.5 Equazioni differenziali scalari lineari a coefficienti costanti

Si tratta di equazioni differenziali lineari a coefficienti costanti e di grado n , ovvero esprimibili nella forma:

$$y^{(n)} + a_1 y^{(n-1)} + \cdots + a_n y = s(t), \quad (2.18)$$

particolare della equazione completa. Entrambi questi addendi possono valutarsi senza calcolare esplicitamente e^{tA} , facendo uso dei seguenti risultati:

Teorema 2.5.1 Ogni soluzione dell'equazione omogenea associata alla (2.18) è combinazione lineare delle n funzioni:

$$e^{\lambda_i t}, \quad t e^{\lambda_i t}, \dots, \quad t^{m_i-1} e^{\lambda_i t} \quad i = 1, \dots, k \leq n \quad 1 \leq m_i \leq n$$

dove $\{\lambda_i\}_{i=1}^k$ sono le k radici distinte, eventualmente complesse, del polinomio caratteristico $P(\lambda)$ (2.19), mentre $1 \leq m_i \leq n$ è la molteplicità della radice i -esima². Nel caso in cui la radice $\lambda_j = a_j + j b_j$ sia complessa, la coppia di soluzioni complesse coniugate:

$$t^q e^{\lambda_j t}, \quad t^q e^{\bar{\lambda}_j t} \quad 0 \leq q \leq m_j - 1$$

viene sostituita dalle corrispondenti soluzioni reali:

$$t^q e^{a_j t} \cos(b_j t), \quad t^q e^{a_j t} \sin(b_j t) \quad 0 \leq q \leq m_j - 1.$$

Teorema 2.5.2 Sia $s(t) = p(t)e^{\alpha t}$, con $p(t)$ polinomio di grado k in t e $\alpha \in \mathbb{C}$, la forma del termine forzante dell'equazione differenziale (2.18). Allora la (2.18) ammette una soluzione particolare del tipo:

$$y_0(t) = q(t)t^m e^{\alpha t},$$

dove $q(t)$ è un polinomio di grado k e $m \neq 0$ solo se α è una radice del polinomio caratteristico (2.19) di molteplicità m .

In particolare, qualora $s(t)$ si possa esprimere come sovrapposizione di termini forzanti del tipo considerato nel Teorema 2.5.2, ovvero:

$$s(t) = p_1(t)e^{\alpha_1 t} + \dots + p_s(t)e^{\alpha_s t},$$

allora una soluzione particolare dell'equazione sarà una sovrapposizione di funzioni del tipo:

$$y_0(t) = \sum_{r=1}^s q_r(t)t_r^{m_r} e^{\alpha_r t},$$

essendo $q_r(t)$ un polinomio dello stesso grado di $p_r(t)$ e m_r la (eventuale) molteplicità di α_r qualora questa sia radice del polinomio caratteristico (2.19).

²Naturalmente, per il teorema fondamentale dell'algebra $\sum_i m_i = n$.

Esempio 2.5.1

Determinare la soluzione di:

$$\begin{cases} x'' - 2x' - 8x = 0 \\ x(1) = 1, \quad x'(1) = 0 \end{cases} .$$

Si tratta di una equazione differenziale del secondo ordine a coefficienti costanti omogenea, quindi per determinarne l'integrale generale si può fare uso del Teorema 2.5.1. Il polinomio caratteristico (2.19) si scrive:

$$P(\lambda) = \lambda^2 - 2\lambda - 8 = (\lambda - 4)(\lambda + 2),$$

pertanto le due radici sono $\lambda_1 = 4$ e $\lambda_2 = -2$, cui corrisponde la soluzione generale:

$$x(t) = c_1 e^{4t} + c_2 e^{-2t} \quad c_1, c_2 \in \mathbb{R}.$$

Imponendo le condizioni iniziali:

$$x(1) = c_1 e^4 + c_2 e^{-2} = 1 \quad x'(1) = 4c_1 e^4 - 2c_2 e^{-2} = 0$$

si ricava:

$$c_1 = \frac{e^{-4}}{3} \quad c_2 = \frac{2}{3} e^2$$

per cui la soluzione richiesta è:

$$x(t) = \frac{1}{3} e^{4(t-1)} + \frac{2}{3} e^{2(1-t)}.$$

◆

Esempio 2.5.2

Determinare la soluzione generale dell'equazione:

$$y'' - y = e^x + 2x^2.$$

Iniziamo a studiare l'equazione omogenea associata:

$$z'' - z = 0,$$

cui corrisponde il polinomio caratteristico:

$$\lambda^2 - 1 = 0,$$

che ha le due radici distinte $\lambda_1 = 1$ e $\lambda_2 = -1$. Pertanto, la soluzione generale dell'omogenea associata è:

$$z(x) = c_1 e^x + c_2 e^{-x} \quad c_1, c_2 \in \mathbb{R}.$$

Per determinare una soluzione particolare dell'equazione completa, consideriamo separatamente i due addendi del termine forzante, ognuno dei quali è della forma prevista nel Teorema 2.5.2:

1. equazione $y_1'' - y_1 = e^x$. Il Teorema 2.5.2 garantisce che una soluzione particolare è della forma $y_1(x) = a x e^x$, infatti $p(x)$ è di grado 0, mentre $\alpha = 1$ è radice del polinomio caratteristico di molteplicità 1. Sostituendo nell'equazione si ha:

$$2a e^x + a x e^x - a x e^x = e^x,$$

da cui segue $a = 1/2$, per cui: $y_1(x) = x e^x / 2$.

2. equazione $y_2'' - y_2 = 2x^2$. Poiché il termine forzante è un polinomio di secondo grado, la soluzione sarà ancora un polinomio di secondo grado: $y_2(x) = a_1 x^2 + a_2 x + a_3$, che sostituita nell'equazione conduce a:

$$2a_1 - a_1 x^2 - a_2 x - a_3 = 2x^2 \implies \begin{cases} -a_1 = 2 \\ a_2 = 0 \\ 2a_1 - a_3 = 0 \end{cases},$$

da cui $a_1 = -2$, $a_2 = 0$ e $a_3 = -4$. Infine: $y_2(x) = -2x^2 - 4$.

La soluzione particolare richiesta, pertanto, vale:

$$y_0(x) = y_1(x) + y_2(x) = \frac{x}{2} e^x - 2x^2 - 4$$

e quindi:

$$y(t) = z(t) + y_0(t) = c_1 e^x + c_2 e^{-x} + \frac{x}{2} e^x - 2x^2 - 4$$

è la soluzione generale dell'equazione completa. ◆

Esempio 2.5.3

Risolvere il problema di Cauchy:

$$\begin{cases} y'' + 2y' + 3y = 0 \\ y(0) = 1, \quad y'(0) = 2 \end{cases}.$$

L'equazione è omogenea, per cui basta studiare il polinomio caratteristico:

$$\lambda^2 + 2\lambda + 3 = 0,$$

che ha le due radici complesse coniugate $\lambda_{1,2} = -1 \pm j\sqrt{2}$ di molteplicità 1. Pertanto, per il Teorema 2.5.1 la soluzione generale dell'equazione differenziale è:

$$y(t) = c_1 e^{-t} \cos(\sqrt{2}t) + c_2 e^{-t} \sin(\sqrt{2}t) \quad c_1, c_2 \in \mathbb{R}.$$

Le due costanti si determinano imponendo le condizioni iniziali:

$$\begin{cases} y(0) = c_1 = 1, \\ y'(0) = (c_2\sqrt{2} - c_1) = 2 \end{cases} ,$$

infatti $y'(t) = (c_2\sqrt{2} - c_1) \cos(\sqrt{2}t) \exp(-t) - (c_2 + c_1\sqrt{2}) \sin(\sqrt{2}t) \exp(-t)$. Si ha così

$$c_1 = 1 \quad c_2 = \frac{3}{\sqrt{2}},$$

e quindi la soluzione richiesta è:

$$y(t) = e^{-t} \cos(\sqrt{2}t) + \frac{3}{\sqrt{2}} e^{-t} \sin(\sqrt{2}t).$$



3

Integrali multipli

Obiettivi: Integrali in \mathbb{R}^2 : integrazione per orizzontali e per verticali. Integrali in \mathbb{R}^3 : integrazione per fili e per strati. Cambiamenti di variabile: coordinate polari nel piano, coordinate cilindriche e sferiche nello spazio.

3.1 Integrali nel piano

Data una funzione di due variabili $f: \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}$, si può definirne l'integrale su un compatto $A \subset \Omega$ estendendo la trattazione svolta nel caso delle funzioni di una sola variabile reale. In particolare, si può iniziare a definire l'integrale di una funzione costante a tratti su un rettangolo $A = [a,b] \times [c,d] \subset \Omega$, e poi estendere tramite un passaggio al limite la definizione ad una qualunque funzione continua e ad una vasta classe di domini di integrazione A . L'*integrale doppio* di $f(x,y)$ sull'insieme A viene indicato con:

$$\int_A f(x,y) \, dx dy.$$

Definizione 3.1.1 Una funzione di due variabili per la quale esista l'integrale doppio su un insieme A viene detta *integrabile*.

Nel seguito, ci occuperemo solamente di integrali doppi su insiemi limitati.

Definizione 3.1.2 Dato un insieme limitato $A \subset \mathbb{R}^2$, si definisce *area* di A l'integrale doppio:

$$m(A) = \int_A 1 \, dx dy.$$

L'integrale doppio gode di proprietà analoghe a quello di una funzione di variabile reale. Si consideri, ad esempio, un dominio di integrazione A costituito dall'unione di un numero finito di sottoinsiemi A_i . Allora vale:

Teorema 3.1.1 Sia f una funzione continua di due variabili integrabile sull'insieme A , che si possa esprimere come:

$$A = \bigcup_{k=1}^n A_k,$$

dove i sottoinsiemi A_k si intersechino a due a due al più lungo le rispettive frontiere. Allora:

$$\int_A f(x,y) \, dx dy = \sum_{k=1}^n \int_{A_k} f(x,y) \, dx dy.$$

Teorema 3.1.2 Siano f, g funzioni continue di due variabili integrabili sull'insieme A , e siano $\alpha, \beta \in \mathbb{R}$. Allora valgono le seguenti proprietà:

1. *linearità*:

$$\int_A [\alpha f(x,y) + \beta g(x,y)] \, dx dy = \alpha \int_A f(x,y) \, dx dy + \beta \int_A g(x,y) \, dx dy;$$

2. *monotonia*:

$$f \geq g \text{ in } A \implies \int_A f(x,y) \, dx dy \geq \int_A g(x,y) \, dx dy.$$

In particolare, l'integrale di una funzione non negativa è anch'esso non negativo;

3. *annullamento*:

$$m(A) = 0 \implies \int_A f(x,y) \, dx dy = 0$$

$$m(A) > 0, f(x,y) \geq 0 \text{ in } A, \int_A f(x,y) \, dx dy = 0 \implies f(x,y) = 0 \text{ in } A;$$

4. se $\Omega \subset \mathbb{R}^2$ è un insieme aperto su cui la funzione f è continua, e se per ogni compatto $A \subset \Omega$ si ha:

$$\int_A f(x,y) \, dx dy = 0$$

allora $f(x,y) = 0$ in tutto Ω .

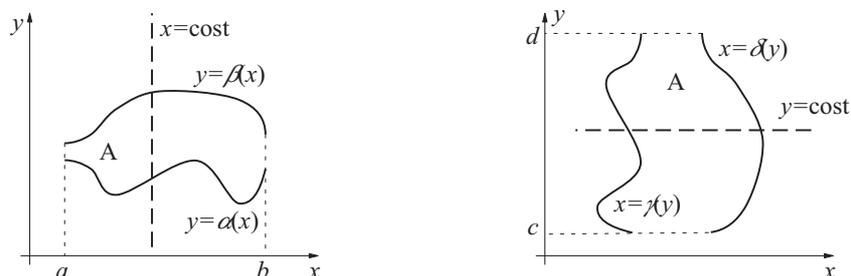


Figura 3.1. Interpretazione geometrica di un insieme verticalmente (a sinistra) e orizzontalmente (a destra) convesso nel piano.

È possibile ricondurre il calcolo di un integrale in \mathbb{R}^2 al calcolo di due integrali semplici successivi, almeno per un particolare insieme di domini di integrazione:

Definizione 3.1.3 Si consideri un insieme compatto $A \subset \mathbb{R}^2$. Esso viene detto *verticalmente convesso* se può essere espresso nella forma:

$$A = \{(x,y) \in \mathbb{R}^2 : a \leq x \leq b, \alpha(x) \leq y \leq \beta(x)\}.$$

Si dice, invece, *orizzontalmente convesso* se può essere espresso nella forma:

$$A = \{(x,y) \in \mathbb{R}^2 : c \leq y \leq d, \gamma(y) \leq x \leq \delta(y)\}.$$

Da un punto di vista geometrico, questa definizione può essere interpretata nel modo seguente: un insieme A è verticalmente (risp. orizzontalmente) convesso se la sua intersezione con ogni retta $x = \text{cost}$ (risp. $y = \text{cost}$) è un segmento o l'insieme vuoto (si veda la Fig. 3.1)

Teorema 3.1.3 Sia $f: \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ una funzione continua e sia $A \subset \mathbb{R}^2$ un insieme verticalmente convesso. Allora (*integrazione per verticali*):

$$\int_A f(x,y) \, dx dy = \int_a^b \left\{ \int_{\alpha(x)}^{\beta(x)} f(x,y) \, dy \right\} dx.$$

Se invece A è orizzontalmente convesso, si ha (*integrazione per orizzontali*):

$$\int_A f(x,y) \, dx dy = \int_c^d \left\{ \int_{\gamma(y)}^{\delta(y)} f(x,y) \, dx \right\} dy.$$

Ciò significa che, per domini di integrazione che siano insiemi orizzontalmente e/o verticalmente convessi, è possibile valutare l'integrale doppio calcolando in sequenza due integrali semplici. Ad esempio, se A è verticalmente convesso, si calcola prima l'integrale rispetto alla variabile y , assumendo x costante e quindi ottenendo una funzione di x :

$$f_1(x) = \int_{\alpha(x)}^{\beta(x)} f(x,y) \, dy,$$

e poi si completa il calcolo valutando l'integrale rispetto ad x di $f_1(x)$:

$$\int_A f(x,y) \, dx dy = \int_a^b f_1(x) \, dx.$$

Esempio 3.1.1

Determinare l'area dell'insieme $A = [1,3] \times [0,1]$. Il dominio A è sia verticalmente sia orizzontalmente convesso, per cui:

$$m(A) = \int_A 1 \, dx dy = \int_0^1 \left(\int_1^3 1 \, dx \right) dy = \int_0^1 2 \, dy = 2.$$

◆

Esempio 3.1.2

Integrare la funzione $f(x,y) = 2xy$ sul dominio $A = \{(x,y) \in \mathbb{R}^2 : 0 \leq x \leq 1, x^2 \leq y \leq x+1\}$. Si tratta di un dominio sia orizzontalmente sia verticalmente convesso, ma vista la forma esplicita di A è più conveniente integrare per verticali:

$$\int_A f(x,y) \, dx dy = \int_0^1 \left(\int_{x^2}^{x+1} 2xy \, dy \right) dx,$$

dove:

$$f_1(x) = \int_{x^2}^{x+1} 2xy \, dy = xy^2 \Big|_{x^2}^{x+1} = x[(x+1)^2 - x^4] = x^3 + 2x^2 + x - x^5.$$

Infine si ha:

$$\begin{aligned} \int_A f(x,y) \, dx dy &= \int_0^1 f_1(x) \, dx = \int_0^1 (x^3 + 2x^2 + x - x^5) \, dx \\ &= \left[\frac{x^4}{4} + \frac{2x^3}{3} + \frac{x^2}{2} - \frac{x^6}{6} \right]_0^1 = \frac{15}{12} = \frac{5}{4}. \end{aligned}$$

◆

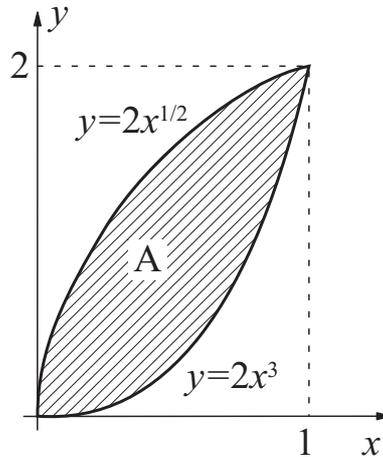


Figura 3.2. Rappresentazione grafica del dominio di integrazione A per l'Esempio 3.1.3.

Esempio 3.1.3

Integrare la funzione $f(x, y) = x + y$ sull'insieme $A = \{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq 1, 2x^3 \leq y \leq 2\sqrt{x}\}$. Anche in questo caso il dominio (si veda la Fig. 3.2) è sia verticalmente sia orizzontalmente convesso, infatti:

$$\begin{aligned} A &= \{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq 1, \quad 2x^3 \leq y \leq 2\sqrt{x}\} \\ &= \{(x, y) \in \mathbb{R}^2 : 0 \leq y \leq 2, \quad (y/2)^2 \leq x \leq (y/2)^{1/3}\}. \end{aligned}$$

Si può pertanto integrare per verticali, ottenendo:

$$f_1(x) = \int_{2x^3}^{2\sqrt{x}} (x + y) \, dy = \frac{(x + y)^2}{2} \Big|_{2x^3}^{2\sqrt{x}} = 2x + 2x^{3/2} - 2x^4 - 2x^6,$$

per cui:

$$\int_A f(x, y) \, dx dy = \int_0^1 f_1(x) \, dx = \left[x^2 + \frac{4}{5}x^{5/2} - \frac{2}{5}x^5 - \frac{2}{7}x^7 \right]_0^1 = \frac{39}{35}.$$

Invece, integrando per orizzontali si ha:

$$f_2(y) = \int_{(y/2)^2}^{(y/2)^{1/3}} (x + y) \, dx = \frac{(x + y)^2}{2} \Big|_{(y/2)^2}^{(y/2)^{1/3}} = -\frac{y^4}{32} - \frac{y^3}{4} + \frac{y^{2/3}}{2^{5/3}} + \frac{y^{4/3}}{2^{1/3}},$$

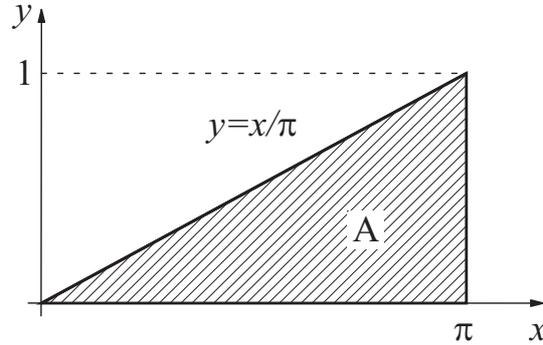


Figura 3.3. Rappresentazione grafica del dominio di integrazione A per l'Esempio 3.1.4.

e quindi:

$$\int_A f(x,y) \, dx dy = \int_0^2 f_2(y) \, dy = \left[-\frac{y^5}{5 \cdot 32} - \frac{y^4}{16} + \frac{3y^{5/3}}{5 \cdot 2^{5/3}} + \frac{3y^{7/3}}{7 \cdot 2^{1/3}} \right]_0^2 = \frac{39}{35}.$$

Naturalmente, con entrambe le tecniche si ottiene lo stesso risultato. \blacklozenge

Esempio 3.1.4

Integrare la funzione $f(x,y) = (\sin x)/x$ sul triangolo di vertici $(0,0)$, $(\pi,0)$ e $(\pi,1)$. Sebbene il dominio sia verticalmente e orizzontalmente convesso (si veda la Fig. 3.2), non è possibile calcolare esplicitamente l'integrale per orizzontali, infatti la funzione $(\sin x)/x$ non è integrabile elementarmente rispetto a x . Per verticali, invece, si ha:

$$A = \{(x,y) \in \mathbb{R}^2 : 0 \leq x \leq \pi, \quad 0 \leq y \leq x/\pi\},$$

e quindi:

$$f_1(x) = \int_0^{x/\pi} \frac{\sin x}{x} \, dy = \frac{\sin x}{x} \frac{x}{\pi} = \frac{\sin x}{\pi}.$$

Infine:

$$\int_A f(x,y) \, dx dy = \int_0^\pi \frac{\sin x}{\pi} \, dx = -\frac{\cos x}{\pi} \Big|_0^\pi = \frac{2}{\pi}.$$

\blacklozenge

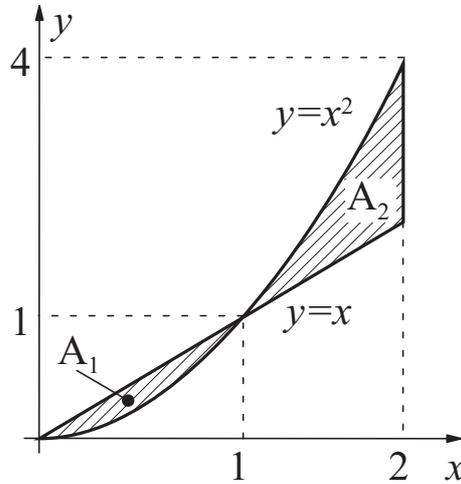


Figura 3.4. Rappresentazione grafica del dominio di integrazione A per l'Esempio 3.1.5.

Esempio 3.1.5

Calcolare l'integrale doppio della funzione $f(x,y) = x + 2y$ sul dominio $A = A_1 \cup A_2$, dove:

$$A_1 = \{(x,y) \in \mathbb{R}^2 : 0 \leq x \leq 1, \quad x^2 \leq y \leq x\}$$

$$A_2 = \{(x,y) \in \mathbb{R}^2 : 1 \leq x \leq 2, \quad x \leq y \leq x^2\}.$$

Le due componenti del dominio A sono rappresentate in Fig. 3.4: poiché la loro intersezione si riduce ad un punto delle relative frontiere, possiamo decomporre l'integrale secondo:

$$\int_A f(x,y) \, dx dy = \int_{A_1} f(x,y) \, dx dy + \int_{A_2} f(x,y) \, dx dy,$$

dove entrambi gli integrali parziali possono essere valutati integrando per verticali:

$$\int_{A_1} f(x,y) \, dx dy = \int_0^1 f_{11}(x) dx \quad \int_{A_2} f(x,y) \, dx dy = \int_1^2 f_{12}(x) dx,$$

dove:

$$f_{11}(x) = \int_{x^2}^x (x + 2y) \, dy = [xy + y^2]_{x^2}^x = 2x^2 - x^3 - x^4,$$

$$f_{12}(x) = \int_x^{x^2} (x + 2y) \, dy = [xy + y^2]_x^{x^2} = -2x^2 + x^3 + x^4.$$

Integrando lungo x :

$$\int_0^1 f_{11}(x) \, dx = \int_0^1 (2x^2 - x^3 - x^4) \, dx = \left[\frac{2}{3}x^3 - \frac{x^4}{4} - \frac{x^5}{5} \right]_0^1 = \frac{13}{60},$$

$$\int_1^2 f_{12}(x) \, dx = \int_1^2 (-2x^2 + x^3 + x^4) \, dx = \left[-\frac{2}{3}x^3 + \frac{x^4}{4} + \frac{x^5}{5} \right]_1^2 = \frac{317}{60}.$$

Infine, l'integrale richiesto vale:

$$\int_A f(x,y) \, dx dy = \frac{13}{60} + \frac{317}{60} = \frac{11}{2}.$$

◆

Da un punto di vista geometrico, l'integrale doppio di una funzione $f(x,y)$ su un dominio di integrazione A rappresenta il volume del solido sotteso alla superficie $z = f(x,y)$.

Esempio 3.1.6

Calcolare il volume V del solido compreso tra il rettangolo $A = [0,1] \times [0,2]$ e il piano $z = f(x,y) = x + y$. Grazie alla definizione di integrale doppio, si ha:

$$V = \int_A f(x,y) \, dx dy = \int_0^2 \left[\int_0^1 (x + y) \, dx \right] dy = \int_0^2 \left(\frac{1}{2} + y \right) dy = 3.$$

◆

L'integrale doppio consente anche di definire alcuni concetti particolarmente importanti in fisica. Si consideri un insieme $A \subset \mathbb{R}^2$ caratterizzato da una densità superficiale $\rho(x,y)$ (ovvero, da una massa ρ per unità di superficie). Si definiscono:

1. il *baricentro* di A come quel punto $\bar{\mathbf{x}} \in \mathbb{R}^2$ di coordinate:

$$\bar{x} = \frac{1}{M} \int_A x \rho(x,y) \, dx dy \quad \bar{y} = \frac{1}{M} \int_A y \rho(x,y) \, dx dy, \quad (3.1)$$

dove M indica la massa totale associata ad A :

$$M = \int_A \rho(x,y) \, dx dy. \quad (3.2)$$

In particolare, se la densità è uniforme (cioè $\rho = \text{cost}$) le due relazioni precedenti si riducono a:

$$\bar{x} = \frac{1}{m(A)} \int_A x \, dx dy \quad \bar{y} = \frac{1}{m(A)} \int_A y \, dx dy, \quad (3.3)$$

e il baricentro viene detto *centroide*;

2. il *momento d'inerzia* di A rispetto all'asse ortogonale al piano (x,y) passante per l'origine come la quantità:

$$I = \int_A d^2(\mathbf{x}, \mathbf{0}) \rho(\mathbf{x}) \, dx dy, \quad (3.4)$$

dove $d(\mathbf{x}, \mathbf{0})$ è la distanza del punto $\mathbf{x} = (x,y)$ dall'origine, ovvero:

$$d^2(\mathbf{x}, \mathbf{0}) = x^2 + y^2.$$

In particolare, se $\rho = \text{cost}$ si ha:

$$I = \frac{M}{m(A)} \int_A d^2(\mathbf{x}, \mathbf{0}) \, dx dy. \quad (3.5)$$

Esempio 3.1.7

Calcolare il baricentro del dominio triangolare A utilizzato nell'Esempio 3.1.4, assumendo che la densità superficiale sia costante. L'area del triangolo può essere valutata immediatamente in $m(A) = \pi/2$, mentre per il baricentro si può fare uso della (3.3):

$$\begin{aligned} \bar{x} &= \frac{1}{m(A)} \int_A x \, dx dy = \frac{2}{\pi} \int_0^\pi \left[\int_0^{x/\pi} x dy \right] dx = \frac{2}{\pi^2} \int_0^\pi x^2 dx = \frac{2\pi}{3} \\ \bar{y} &= \frac{1}{m(A)} \int_A y \, dx dy = \frac{2}{\pi} \int_0^\pi \left[\int_0^{x/\pi} y dy \right] dx = \frac{1}{\pi^3} \int_0^\pi x^2 dx = \frac{1}{3}. \end{aligned}$$



3.2 Integrali nello spazio

Grazie ad un procedimento analogo a quello delineato nella Sez. 3.1, è possibile definire l'integrale di una funzione continua di più variabili $f: \Omega \subset \mathbb{R}^3 \rightarrow \mathbb{R}$ su un compatto $A \subset \Omega$. Si parla, in questo caso, di *integrale triplo*:

$$\int_A f(x,y,z) \, dx dy dz.$$

Definizione 3.2.1 Una funzione di tre variabili per la quale esista l'integrale triplo su un insieme $A \subset \mathbb{R}^3$ viene detta *integrabile*.

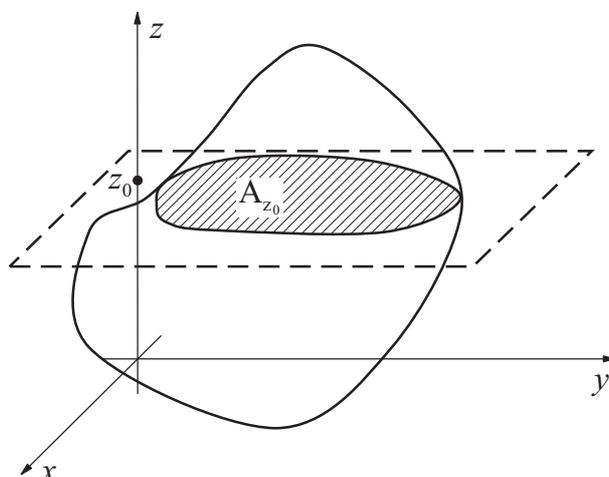


Figura 3.5. Rappresentazione grafica dell'intersezione A_{z_0} del piano $z = z_0$ con un dominio di integrazione A limitato in \mathbb{R}^3 .

Anche in questo caso, ci occuperemo solamente di integrali tripli su insiemi limitati.

Definizione 3.2.2 Dato un insieme limitato $A \subset \mathbb{R}^3$, si definisce *volume* di A l'integrale triplo:

$$m(A) = \int_A 1 \, dx dy dz.$$

Si noti come anche per l'integrale triplo valgano i risultati enunciati per l'integrale doppio relativi all'additività e alla linearità, ovvero i Teoremi 3.1.1 e 3.1.2 opportunamente riformulati.

Il calcolo degli integrali tripli può essere ricondotto alla valutazione di integrali semplici e doppi, in funzione della forma del dominio di integrazione A . In particolare, per ogni $z_0 \in \mathbb{R}$ si definisce l'insieme:

$$A_{z_0} = \{(x, y) \in \mathbb{R}^2 : (x, y, z_0) \in A\} \subset \mathbb{R}^2.$$

In altri termini, A_{z_0} è l'intersezione di A con il piano $z = z_0$ (si veda la Fig. 3.5). Poiché A è limitato, A_{z_0} sarà vuoto se z_0 è al di fuori di un certo intervallo $I \subset \mathbb{R}$.

Se l'insieme A è sufficientemente regolare, vale il seguente

Teorema 3.2.1 Sia $f: \Omega \subset \mathbb{R}^3 \rightarrow \mathbb{R}$ una funzione continua e sia $A \subset \Omega$. Se la funzione di due variabili $f(x, y, z_0)$ è integrabile in $A_{z_0} \forall z_0 \in I$, allora (*integrazione*

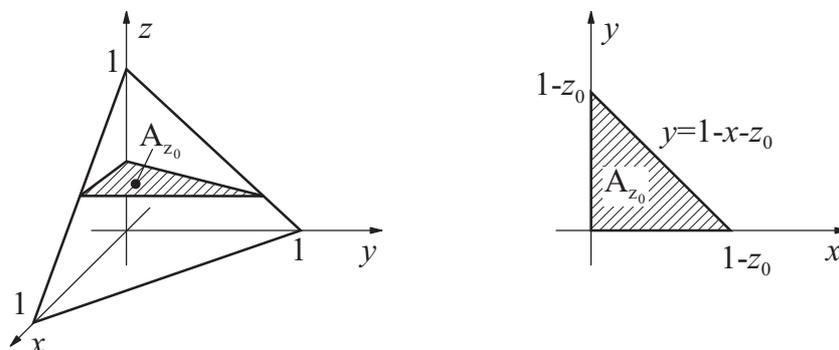


Figura 3.6. Rappresentazione grafica del dominio di integrazione A (a sinistra) e della sezione A_{z_0} (a destra) per l'Esempio 3.2.1.

per strati paralleli al piano (x,y)):

$$\int_A f(x,y,z) \, dx dy dz = \int_I \left\{ \int_{A_z} f(x,y,z) \, dx dy \right\} dz.$$

L'altra possibile tecnica presuppone che l'insieme A sia convesso nella direzione dell'asse z , ovvero che si possa esprimere:

$$A = \{(x,y,z) \in \mathbb{R}^3 : (x,y) \in D, \quad \alpha(x,y) \leq z \leq \beta(x,y)\}$$

dove α e β sono due funzioni di due variabili continue sul compatto D . In questo caso:

Teorema 3.2.2 Sia $f: \Omega \subset \mathbb{R}^3 \rightarrow \mathbb{R}$ una funzione continua e sia $A \subset \Omega$ convesso rispetto all'asse z . Allora (*integrazione per fili* paralleli all'asse z):

$$\int_A f(x,y,z) \, dx dy dz = \int_D \left\{ \int_{\alpha(x,y)}^{\beta(x,y)} f(x,y,z) \, dz \right\} dx dy.$$

Nel formulare gli ultimi due Teoremi, abbiamo privilegiato l'asse z : naturalmente, è possibile dedurre formule di riduzione analoghe prendendo a riferimento gli altri due assi dello spazio, ovvero ricavare formule per la integrazione per strati paralleli ai piani (x,z) e (y,z) , e formule per la integrazione per fili paralleli agli assi x e y .

Esempio 3.2.1

Integrare la funzione $f(x,y,z) = x$ sull'insieme A contenuto nel primo ottante e delimitato dal piano di equazione $x+y+z = 1$. Il dominio di integrazione, rappresentato

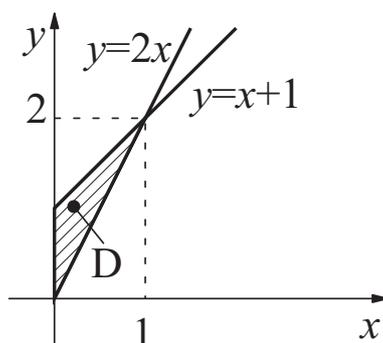


Figura 3.7. Rappresentazione grafica del dominio di integrazione D per l'Esempio 3.2.2.

nella parte sinistra della Fig. 3.6, può essere decomposto per strati paralleli all'asse z . In particolare, $A_{z_0} \neq \emptyset$ solo per $z_0 \in I = [0, 1]$, mentre per $z_0 \in I$ la sezione A_{z_0} è rappresentata nella parte destra della Fig. 3.6. L'integrale si può quindi calcolare per strati:

$$\int_A f(x, y, z) \, dx dy dz = \int_0^1 \left\{ \int_{A_z} f(x, y, z) \, dx dy \right\} dz,$$

dove l'integrale doppio può essere calcolato, ad esempio, per verticali:

$$\int_{A_z} f(x, y, z) \, dx dy = \int_0^{1-z} \left(\int_0^{1-x-z} x \, dy \right) dx = \int_0^{1-z} [(1-z)x - x^2] dx = \frac{1}{6}(1-z)^3.$$

Infine si ha:

$$\int_A f(x, y, z) \, dx dy dz = \int_0^1 \frac{1}{6}(1-z)^3 \, dz = -\frac{(1-z)^4}{24} \Big|_0^1 = \frac{1}{24}.$$

◆

Esempio 3.2.2

Calcolare l'integrale della funzione $f(x, y, z) = x + y + z$ sul dominio A :

$$A = \{(x, y, z) \in \mathbb{R}^3 : (x, y) \in D, \quad 0 \leq z \leq x + y\},$$

essendo $D = \{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq 1, \quad 2x \leq y \leq x + 1\}$. Il dominio A è convesso rispetto all'asse z , mentre D (rappresentato in Fig. 3.7) è verticalmente convesso nel piano (naturalmente, è anche orizzontalmente convesso nel piano, ma nella prima

interpretazione l'espressione analitica di D è più semplice). Allora, integrando per fili:

$$\int_A f(x,y,z) \, dx dy dz = \int_D \left\{ \int_0^{x+y} f(x,y,z) \, dz \right\} \, dx dy,$$

dove:

$$\int_0^{x+y} f(x,y,z) \, dz = \int_0^{x+y} (x+y+z) \, dz = \frac{3}{2}(x+y)^2.$$

Integrando per verticali si ha poi:

$$\int_D \frac{3}{2}(x+y)^2 \, dx dy = \frac{3}{2} \int_0^1 \left[\int_{2x}^{x+1} (x+y)^2 \, dy \right] \, dx,$$

dove:

$$\frac{3}{2} \int_{2x}^{x+1} (x+y)^2 \, dy = \frac{(2x+1)^3}{2} - \frac{27x^3}{2}.$$

Infine:

$$\int_0^1 \left[\frac{(2x+1)^3}{2} - \frac{27x^3}{2} \right] \, dx = \left[\frac{(2x+1)^4}{16} - \frac{27x^4}{8} \right]_0^1 = \frac{13}{8}.$$

◆

Anche l'integrale triplo permette di definire baricentro e momento di inerzia associati ad un volume $A \subset \mathbb{R}^3$ caratterizzato da una densità $\rho(x,y,z)$ (ovvero, da una massa ρ per unità di volume). In particolare, il *baricentro* di A è definito come quel punto $\bar{\mathbf{x}} \in \mathbb{R}^3$ di coordinate:

$$\begin{aligned} \bar{x} &= \frac{1}{M} \int_A x \rho(x,y,z) \, dx dy dz \\ \bar{y} &= \frac{1}{M} \int_A y \rho(x,y,z) \, dx dy dz \\ \bar{z} &= \frac{1}{M} \int_A z \rho(x,y,z) \, dx dy dz, \end{aligned}$$

dove M indica la massa totale associata ad A :

$$M = \int_A \rho(x,y,z) \, dx dy dz.$$

In particolare, se la densità è uniforme (cioè $\rho = \text{cost}$) le relazioni precedenti si riducono a:

$$\bar{x} = \frac{1}{m(A)} \int_A x \, dx dy dz \quad \bar{y} = \frac{1}{m(A)} \int_A y \, dx dy dz \quad \bar{z} = \frac{1}{m(A)} \int_A z \, dx dy dz.$$

3.3 Cambiamenti di variabile

Nel calcolo degli integrali semplici, spesso si fa uso del cambiamento di variabile con lo scopo di modificare la forma della funzione integranda per rendere più semplice la valutazione della primitiva. Anche nel caso degli integrali multipli è possibile fare uso di cambiamenti di variabile, ma in questo caso lo scopo, di solito, è semplificare il dominio di integrazione.

Con cambiamento di variabile, si intende una trasformazione biunivoca $\Phi: \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ che trasformi il dominio di integrazione originale A in un altro dominio $A' = \Phi^{-1}(A)$. Posto $\mathbf{x} = \Phi(\mathbf{u}) \in A$, immagine del punto $\mathbf{u} \in A'$, si definisce la *matrice Jacobiana* $\mathbf{J}_\Phi(\mathbf{u})$ della trasformazione $\Phi(\mathbf{u}) = (\Phi^{(1)}(\mathbf{u}), \dots, \Phi^{(n)}(\mathbf{u}))$ la collezione di derivate parziali:

$$\mathbf{J}_\Phi(\mathbf{u}) = \begin{bmatrix} \Phi_{u_1}^{(1)}(\mathbf{u}) & \Phi_{u_2}^{(1)}(\mathbf{u}) & \dots & \Phi_{u_n}^{(1)}(\mathbf{u}) \\ \Phi_{u_1}^{(2)}(\mathbf{u}) & \Phi_{u_2}^{(2)}(\mathbf{u}) & \dots & \Phi_{u_n}^{(2)}(\mathbf{u}) \\ \vdots & \vdots & \ddots & \vdots \\ \Phi_{u_1}^{(n)}(\mathbf{u}) & \Phi_{u_2}^{(n)}(\mathbf{u}) & \dots & \Phi_{u_n}^{(n)}(\mathbf{u}) \end{bmatrix}.$$

Teorema 3.3.1 Sia $\Phi: \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ una funzione di classe $C^1(\Omega)$, con Ω aperto, biunivoca e tale che $\det \{\mathbf{J}_\Phi(\mathbf{u})\}$ sia diverso da zero e limitato $\forall \mathbf{u} \in \Omega$. Siano $A \subset \Phi(\Omega)$ e $A' = \Phi^{-1}(A) \subset \Omega$ insiemi limitati sui quali la funzione di più variabili 1 sia integrabile¹. Allora, se f è una funzione continua su A , si ha la seguente *formula di cambiamento di variabile*:

$$\int_A f(\mathbf{x}) \, d\mathbf{x} = \int_{A'} f(\Phi(\mathbf{u})) \det \{\mathbf{J}_\Phi(\mathbf{u})\} \, d\mathbf{u}.$$

Esempio 3.3.1

Si consideri un insieme $A \subset \mathbb{R}^2$. Nel piano si definisce con *coordinate polari* un sistema di riferimento caratterizzato dalla distanza ρ dall'origine e dall'angolo θ rispetto all'asse x , come indicato in Fig. 3.8:

$$x = \Phi^{(1)}(\rho, \theta) = \rho \cos \theta, \quad y = \Phi^{(2)}(\rho, \theta) = \rho \sin \theta.$$

La matrice Jacobiana associata alla trasformazione ha componenti:

$$\frac{\partial x}{\partial \rho} = \cos \theta, \quad \frac{\partial x}{\partial \theta} = -\rho \sin \theta, \quad \frac{\partial y}{\partial \rho} = \sin \theta, \quad \frac{\partial y}{\partial \theta} = \rho \cos \theta,$$

¹Ovvero, si possano valutare $m(A)$ e $m(A')$.

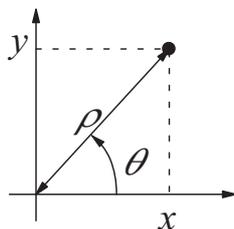


Figura 3.8. Rappresentazione grafica delle coordinate polari nel piano.

per cui:

$$\det \{J_{\Phi}(\rho, \theta)\} = \begin{vmatrix} \cos \theta & -\rho \sin \theta \\ \sin \theta & \rho \cos \theta \end{vmatrix} = \rho \cos^2 \theta + \rho \sin^2 \theta = \rho.$$

Applicando il Teorema 3.3.1 si ha quindi il cambiamento di variabile:

$$dxdy = \rho \, d\rho d\theta,$$

che costituisce l'espressione dell'elemento di superficie in coordinate polari. Per quanto riguarda la funzione integranda, naturalmente si esprime:

$$f(x, y) = f(\rho \cos \theta, \rho \sin \theta).$$



Esempio 3.3.2

Integrare la funzione $f(x, y) = \left| \ln \sqrt{x^2 + y^2} \right|$ sull'insieme A definito dalle disuguaglianze $\epsilon^2 \leq x^2 + y^2 \leq 1$, $0 < \epsilon < 1$.

Il dominio di integrazione A è la corona circolare compresa tra le circonferenze di centro l'origine e raggi ϵ e 1 (Fig. 3.3.2 a sinistra), e quindi può essere rappresentata in coordinate polari dalla regione rettangolare (Fig. 3.3.2 a destra):

$$A = \{(\rho, \theta) : \epsilon \leq \rho \leq 1, \quad 0 \leq \theta \leq 2\pi\}.$$

Nella funzione integranda, il cambiamento di variabile conduce a:

$$f(\rho \cos \theta, \rho \sin \theta) = |\ln \rho| = -\ln \rho \quad \text{essendo } \rho \leq 1.$$

Infine, si ha:

$$\int_A f(x, y) \, dxdy = \int_{A'} f(\rho \cos \theta, \rho \sin \theta) \rho \, d\rho d\theta,$$

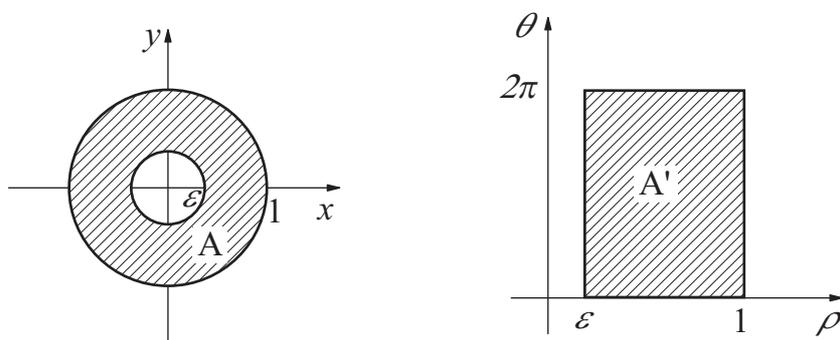


Figura 3.9. Rappresentazione grafica del dominio di integrazione A per l'Esempio 3.3.2.

dove il dominio di integrazione A' è un rettangolo in coordinate polari. Si ha quindi:

$$\int_{A'} f(\rho \cos \theta, \rho \sin \theta) \rho \, d\rho d\theta = - \int_0^{2\pi} \int_{\epsilon}^1 \rho \ln \rho \, d\rho d\theta,$$

dove, integrando per parti:

$$\int \rho \ln \rho \, d\rho = \frac{\rho^2}{2} \ln \rho - \int \frac{\rho^2}{2} \frac{1}{\rho} \, d\rho = \frac{\rho^2}{2} \ln \rho - \frac{\rho^2}{4}.$$

Infine:

$$\int_A f(x,y) \, dx dy = \int_0^{2\pi} \left[\frac{1}{4} + \frac{\epsilon^2}{2} \ln \epsilon - \frac{\epsilon^2}{4} \right] \, d\theta = \pi \left[\frac{1}{2} + \epsilon^2 \ln \epsilon - \frac{\epsilon^2}{2} \right].$$

◆

Esempio 3.3.3

Si consideri un insieme $A \subset \mathbb{R}^3$. Nello spazio, oltre alle coordinate cartesiane utilizzate finora, si possono definire altri due sistemi di riferimento particolarmente utili quando l'insieme A presenta particolari simmetrie:

1. *coordinate cilindriche*, che fanno uso di una coordinata lungo uno degli assi cartesiani, ad esempio l'asse z , e di un sistema di coordinate polari nel piano a questo ortogonale, in questo caso il piano (x,y) . Dalla parte sinistra della Fig. 3.10 si deduce facilmente:

$$x = \Phi^{(1)}(\rho, \theta, t) = \rho \cos \varphi, \quad y = \Phi^{(2)}(\rho, \theta, t) = \rho \sin \varphi, \quad z = \Phi^{(3)}(\rho, \theta, t) = t,$$

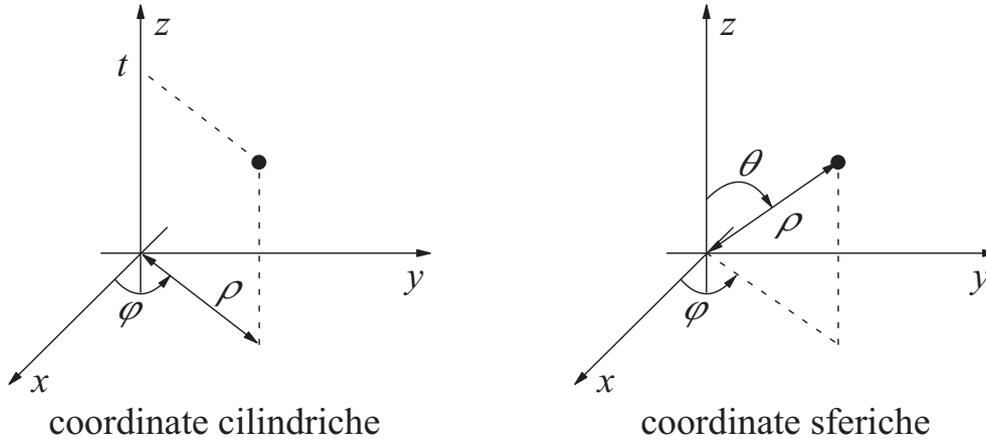


Figura 3.10. Rappresentazione grafica delle coordinate cilindriche (a sinistra) e sferiche (a destra) nello spazio.

pertanto:

$$\det \{J_{\Phi}(\rho, \theta)\} = \begin{vmatrix} \cos \varphi & -\rho \sin \varphi & 0 \\ \sin \varphi & \rho \cos \varphi & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1[\rho \cos^2 \varphi + \rho \sin^2 \varphi] = \rho.$$

L'elemento di volume in coordinate cilindriche, quindi, si esprime:

$$dx dy dz = \rho \, d\rho \, d\varphi \, dt.$$

2. *coordinate sferiche*, che fanno uso di un sistema di riferimento definito dalla distanza dall'origine, e dagli angoli che il vettore forma con gli assi x e z , secondo la rappresentazione nella parte destra della Fig. 3.10. Si ha facilmente:

$$x = \Phi^{(1)}(\rho, \theta, \varphi) = \rho \sin \theta \cos \varphi$$

$$y = \Phi^{(2)}(\rho, \theta, \varphi) = \rho \sin \theta \sin \varphi$$

$$z = \Phi^{(3)}(\rho, \theta, \varphi) = \rho \cos \theta,$$

e quindi:

$$\det \{J_{\Phi}(\rho, \theta)\} = \begin{vmatrix} \sin \theta \cos \varphi & \rho \cos \theta \cos \varphi & -\rho \sin \theta \sin \varphi \\ \sin \theta \sin \varphi & \rho \cos \theta \sin \varphi & \rho \sin \theta \cos \varphi \\ \cos \theta & -\rho \sin \theta & 0 \end{vmatrix} = \rho^2 \sin \theta.$$

L'elemento di volume in coordinate sferiche, quindi, si esprime:

$$dx dy dz = \rho^2 \sin \theta \, d\rho d\theta d\varphi.$$



4

Integrali di linea e di flusso

Obiettivi: Richiami su curve e superfici. Integrale di linea: campi conservativi, Teorema di Green. Integrali di flusso: Teorema di Stokes, Teorema di Gauss.

4.1 Richiami su curve e superfici

L'introduzione dei principali teoremi della teoria dei campi vettoriali richiede la definizione dei concetti fondamentali su curve e superfici, già brevemente accennati nella Sez. 1.1.2.

4.1.1 Curve

Definizione 4.1.1 Una *curva* γ in \mathbb{R}^n è una applicazione continua da $I \subset \mathbb{R}$ in \mathbb{R}^n , dove I è un intervallo, aperto ed eventualmente non limitato. Se, invece, $I = [a, b]$ è un intervallo chiuso e limitato, la restrizione di γ ad $[a, b]$ viene detta *arco di curva*.

Definizione 4.1.2 Una curva $\gamma: I \subset \mathbb{R} \rightarrow \mathbb{R}^n$ si dice di *classe* $C^1(I)$ se la funzione γ è di classe $C^1(I)$. Un arco di curva $\gamma: [a, b] \subset \mathbb{R} \rightarrow \mathbb{R}^n$ è di classe $C^1([a, b])$ se la funzione γ è di classe $C^1([a, b])$.

Definizione 4.1.3 Se γ è un arco di classe $C^1([a, b])$ e $t \in [a, b]$, il vettore delle derivate prime $\gamma'(t)$ è il *vettore tangente* all'arco γ nel punto $\gamma(t)$.

Definizione 4.1.4 Un arco di curva $\gamma: [a, b] \subset \mathbb{R} \rightarrow \mathbb{R}^n$ si dice di *regolare* se è di classe $C^1([a, b])$ e se $\gamma'(t) \neq \mathbf{0} \forall t \in [a, b]$. Si dice che γ è *regolare a tratti* se esiste

un numero finito di punti $a = t_0 < t_1 < t_2 < \dots < t_k = b$ tali che γ sia regolare su ciascun intervallo $[t_{j-1}, t_j]$, $1 \leq j \leq k$.

Dato un arco di curva $\gamma: [a, b] \subset \mathbb{R} \rightarrow \mathbb{R}^n$, la variabile $t \in [a, b]$ viene detto *parametro* dell'arco, mentre l'insieme immagine:

$$\Gamma = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} = \gamma(t) \text{ per qualche } t \in [a, b]\}$$

è il *sostegno* dell'arco. I punti $\gamma(a)$ e $\gamma(b)$ sono gli *estremi* dell'arco, che è *chiuso* se $\gamma(a) = \gamma(b)$. Se l'applicazione γ è iniettiva, cioè se a valori distinti del parametro in $[a, b]$ corrispondono punti diversi nel sostegno, l'arco viene detto *semplice*; nel caso di un arco chiuso, questo è semplice se gli estremi sono gli unici due punti in cui la funzione non è iniettiva.

Esempio 4.1.1

L'applicazione a valori in \mathbb{R}^2 :

$$\gamma(t) = (\cos t, \sin t) \quad t \in [0, 2\pi]$$

è un arco chiuso, semplice e regolare il cui sostegno coincide con la circonferenza di centro l'origine e raggio unitario caratterizzata dall'equazione $x^2 + y^2 = 1$. Invece, se l'intervallo di definizione del parametro diventa $[0, 4\pi]$, si ha ancora un arco regolare e chiuso con lo stesso sostegno, ma non più semplice. \blacklozenge

Dato un arco di curva $\gamma: [a, b] \subset \mathbb{R} \rightarrow \mathbb{R}^n$, se questo è semplice è possibile definire sul suo sostegno Γ una relazione d'ordine. Infatti, scelti due punti $\mathbf{x}_1, \mathbf{x}_2 \in \Gamma$, essendo l'arco semplice questi corrispondono, rispettivamente, ai due valori t_1 e t_2 del parametro: si dice che \mathbf{x}_1 *precede* \mathbf{x}_2 se $t_1 < t_2$: questa relazione d'ordine viene anche detta *senso di percorrenza* dell'arco.

Definizione 4.1.5 Un *cambiamento di parametro* è una funzione α di classe C^1 su un intervallo limitato e chiuso $[c, d]$, a valori in un altro intervallo limitato e chiuso $[a, b]$, suriettiva¹ e con derivata mai nulla in tutto $[c, d]$.

Dalla definizione, segue immediatamente che un cambiamento di parametro è necessariamente una funzione strettamente monotona, crescente o decrescente a seconda che α' sia positiva o negativa in $[c, d]$.

¹Cioè tale che $\alpha([c, d]) = [a, b]$.

Teorema 4.1.1 Sia $\gamma: [a,b] \subset \mathbb{R} \rightarrow \mathbb{R}^n$ un arco di curva regolare e sia α un cambiamento di parametro, con $\alpha' > 0$. Allora l'arco di curva

$$\delta(\tau) = \gamma(\alpha(\tau)): [c,d] \subset \mathbb{R} \rightarrow \mathbb{R}^n$$

è regolare, ed ha lo stesso sostegno e gli stessi estremi di γ . Inoltre:

- se γ è semplice, anche δ è semplice ed individua sul sostegno lo stesso senso di percorrenza di γ ;
- se γ è chiuso, anche δ è chiuso.

Un cambiamento di parametro con derivata negativa, conserva la regolarità e la semplicità di un arco, ma determina una inversione del senso di percorrenza e lo scambio dei due estremi.

Definizione 4.1.6 Due archi di curva regolari:

$$\gamma: [a,b] \subset \mathbb{R} \rightarrow \mathbb{R}^n \quad \text{e} \quad \delta: [c,d] \subset \mathbb{R} \rightarrow \mathbb{R}^n$$

si dicono *equivalenti* se esiste un cambiamento di parametro $\alpha: [c,d] \rightarrow [a,b]$, con α' positivo, tale che:

$$\delta(\tau) = \gamma(\alpha(\tau)).$$

Sulla base del Teorema 4.1.1, due archi equivalenti hanno lo stesso senso di percorrenza, gli stessi estremi (nello stesso ordine) e lo stesso sostegno.

Definizione 4.1.7 Sia $\gamma: [a,b] \subset \mathbb{R} \rightarrow \mathbb{R}^n$ un arco di curva regolare. Il numero:

$$l_\gamma = \int_a^b |\gamma'(t)| \, dt$$

si chiama *lunghezza* di γ .

Teorema 4.1.2 Siano $\gamma: [a,b] \subset \mathbb{R} \rightarrow \mathbb{R}^n$ e $\delta: [c,d] \subset \mathbb{R} \rightarrow \mathbb{R}^n$ due archi di curva regolari equivalenti. Allora:

$$\int_a^b |\gamma'(t)| \, dt = \int_c^d |\delta'(\tau)| \, d\tau.$$

4.1.2 Superfici

Prima di enunciare le principali definizioni relative alle superfici, occorre definire una particolare classe di insiemi aperti:

Definizione 4.1.8 Sia $A \subset \mathbb{R}^n$. A si dice *connesso* se, dati comunque due punti $\mathbf{x}, \mathbf{y} \in A$ esiste un arco di curva di estremi \mathbf{x} e \mathbf{y} il cui sostegno sia contenuto in A .

Essa ammette il caso particolare:

Definizione 4.1.9 Sia $A \subset \mathbb{R}^n$. A si dice *convesso* se, dati comunque due punti $\mathbf{x}, \mathbf{y} \in A$ il segmento di estremi \mathbf{x} e \mathbf{y} è contenuto in A .

Si può, ora, definire una superficie in forma parametrica:

Definizione 4.1.10 Sia $A \subset \mathbb{R}^2$ un aperto connesso. Si chiama *superficie* in \mathbb{R}^n un'applicazione continua $\sigma: A \rightarrow \mathbb{R}^n$. L'insieme immagine:

$$\Sigma = \sigma(A) = \{\sigma(\mathbf{x}) \in \mathbb{R}^n : \mathbf{x} \in A\}$$

si chiama *sostegno* della superficie.

Definizione 4.1.11 Una superficie $\sigma: A \rightarrow \mathbb{R}^n$ si dice *regolare* se la funzione vettoriale $\sigma(x, y) \in C^1(A)$ e se la relativa matrice Jacobiana $J_\sigma(x, y)$ (di dimensione $n \times 2$) ha rango 2 in ogni punto $(x, y) \in A$.

Si noti come la richiesta di rango 2 per la matrice jacobiana corrisponda al fatto che i due vettori

$$\frac{\partial \sigma}{\partial x} \quad \frac{\partial \sigma}{\partial y},$$

che costituiscono le due colonne di $J_\sigma(x, y)$, siano non nulli e linearmente indipendenti in tutto A . Consideriamo, ora, le due restrizioni della superficie passanti per il punto $\mathbf{x}_0 = (x_0, y_0) \in A$:

$$x \rightarrow \sigma(x, y_0) \quad y \rightarrow \sigma(x_0, y),$$

che definiscono, in un intorno di \mathbf{x}_0 , due curve regolari passanti entrambe per il punto $P_0 = \sigma(\mathbf{x}_0) \in \Sigma$. I vettori tangenti a queste due curve nel punto P_0 sono, rispettivamente:

$$\frac{\partial \sigma}{\partial x}(x_0, y_0) \quad \frac{\partial \sigma}{\partial y}(x_0, y_0)$$

che, se la superficie è regolare, sono anche linearmente indipendenti: in questo caso, individuano un (iper)piano in \mathbb{R}^n che costituisce il *piano tangente* alla superficie σ nel punto P_0 . In altri termini, una superficie regolare in A ammette piano tangente in tutti i punti del suo sostegno $\Sigma = \sigma(A)$.

Definizione 4.1.12 Una superficie $\sigma: A \rightarrow \mathbb{R}^n$ si dice *semplice* se la funzione vettoriale $\sigma(x,y)$ è iniettiva.

Anche per le superfici è possibile definire il concetto di cambiamento di parametro e di superfici equivalenti:

Definizione 4.1.13 Siano $A, A' \subset \mathbb{R}^n$ aperti e convessi. Un *cambiamento di coordinate* per superfici è un'applicazione biunivoca $\alpha: A' \rightarrow A$, di classe $C^1(A')$ e tale che la matrice Jacobiana J_α sia non singolare² in ogni punto di A' .

Teorema 4.1.3 Sia $\sigma: A \rightarrow \mathbb{R}^n$ una superficie regolare, e sia $\alpha: A' \rightarrow A$ un cambiamento di coordinate. Allora la funzione vettoriale $\rho(\xi,\eta) = \sigma(\alpha(\xi,\eta))$: $A' \rightarrow \mathbb{R}^n$ definisce una superficie regolare che ha lo stesso sostegno di σ . Inoltre, ρ è semplice se e solo se σ è semplice.

Definizione 4.1.14 Due superfici $\sigma: A \rightarrow \mathbb{R}^n$ e $\rho: A' \rightarrow \mathbb{R}^n$ si dicono *equivalenti* se esiste un cambiamento di coordinate $\alpha: A' \rightarrow A$, con $\det \{J_\alpha(\xi,\eta)\} > 0 \forall (\xi,\eta) \in A'$, tale che $\rho(\xi,\eta) = \sigma(\alpha(\xi,\eta)) \forall (\xi,\eta) \in A'$.

È immediato verificare come due superfici equivalenti abbiano lo stesso sostegno.

Nell'ambito della teoria dei campi vettoriali, di particolare importanza sono le superfici in \mathbb{R}^3 : considereremo, quindi, una superficie $\sigma: A \rightarrow \mathbb{R}^3$ semplice e regolare. Dato un punto $\mathbf{x}_0 \in A$, cui corrisponde il punto $P_0 = \sigma(\mathbf{x}_0) \in \Sigma$, consideriamo nuovamente i due vettori

$$\frac{\partial \sigma}{\partial x}(x_0, y_0) \quad \frac{\partial \sigma}{\partial y}(x_0, y_0),$$

che identificano il piano tangente alla superficie nel punto P_0 . Poiché essi sono linearmente indipendenti, il loro prodotto vettoriale:

$$\mathbf{n}(\mathbf{x}_0) = \frac{\partial \sigma}{\partial x}(x_0, y_0) \wedge \frac{\partial \sigma}{\partial y}(x_0, y_0)$$

²Come noto, una matrice quadrata è non singolare solo se il suo determinante non è nullo.

è un vettore non nullo ortogonale al piano tangente stesso. Per questo motivo, $\mathbf{n}(\mathbf{x}_0)$ viene detto *vettore normale* alla superficie σ nel punto P_0 .

Definizione 4.1.15 Sia $\Sigma \subset \mathbb{R}^3$ il sostegno della superficie semplice e regolare $\sigma: A \rightarrow \mathbb{R}^3$. Sia $\mathbf{x}_0 = (x_0, y_0) \in A$ e sia $P_0 = \sigma(\mathbf{x}_0) \in \Sigma$. Il versore:

$$\hat{\mathbf{n}}(\mathbf{x}_0) = \frac{\mathbf{n}(\mathbf{x}_0)}{|\mathbf{n}(\mathbf{x}_0)|}$$

si chiama *versore normale* alla superficie σ nel punto P_0 .

Consideriamo ancora una superficie semplice e regolare $\sigma: A \rightarrow \mathbb{R}^3$, e sia $K \subset A$ un sottoinsieme compatto la cui frontiera sia il sostegno di un arco chiuso semplice e regolare a tratti. La restrizione $\sigma_0: K \rightarrow \mathbb{R}^3$ di σ a K viene detta *calotta regolare*. L'immagine $\sigma_0(K)$ viene detta *sostegno* della calotta.

Definizione 4.1.16 Sia $\sigma_0: K \rightarrow \mathbb{R}^3$ una calotta regolare. Si chiama *area* della calotta regolare σ_0 il valore dell'integrale di superficie:

$$\mathcal{A}_{\sigma_0} = \int_K |\mathbf{n}(x,y)| \, dx dy.$$

Teorema 4.1.4 Siano $\sigma_0: K \rightarrow \mathbb{R}^3$ e $\rho_0: K' \rightarrow \mathbb{R}^3$ due calotte equivalenti. Allora:

$$\mathcal{A}_{\sigma_0} = \mathcal{A}_{\rho_0}.$$

4.2 Integrali di linea

Consideriamo un campo vettoriale $\mathbf{F}: A \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$, definito e continuo in un aperto A , e un arco di curva regolare $\gamma: [a,b] \subset \mathbb{R} \rightarrow \mathbb{R}^n$ il cui sostegno sia contenuto in A . Possiamo definire:

Definizione 4.2.1 Sia $\gamma: [a,b] \subset \mathbb{R} \rightarrow \mathbb{R}^n$ un arco di curva regolare, e sia $\mathbf{F}: A \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ un campo vettoriale continuo su un aperto A contenente il sostegno di γ . Si definisce l'*integrale di linea* di \mathbf{F} lungo l'arco γ come:

$$\int_{\gamma} \mathbf{F}(\mathbf{x}) \cdot d\mathbf{x} = \int_a^b \mathbf{F}(\gamma(t)) \cdot \gamma'(t) \, dt.$$

L'integrale di linea lungo un arco chiuso viene anche detto *circuitazione*.

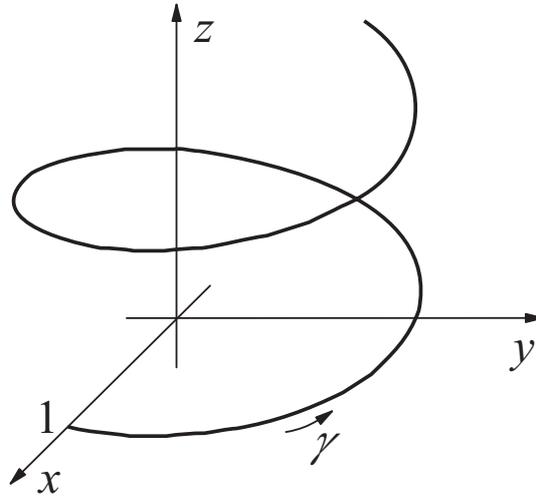


Figura 4.1. Rappresentazione grafica della curva γ utilizzata nell'Esempio 4.2.1. La freccia indica il senso di percorrenza della curva con la parametrizzazione scelta.

Esempio 4.2.1

Si consideri il campo vettoriale:

$$\mathbf{F}(x,y,z) = \left(\frac{x}{x^2 + y^2}, \frac{y}{x^2 + y^2}, -1 \right)$$

definito nell'aperto A costituito da \mathbb{R}^3 privato dell'asse z . Sia poi γ l'arco di elica definito da (Fig. 4.1):

$$\gamma(t) = (\cos t, \sin t, t) \quad t \in [0, 3\pi].$$

Per calcolare l'integrale di linea di \mathbf{F} lungo γ occorre fare uso delle espressioni (infatti $\cos^2 t + \sin^2 t = 1$):

$$\begin{aligned} \mathbf{F}(\gamma(t)) &= (\cos t, \sin t, -1), \\ \gamma'(t) &= (-\sin t, \cos t, 1), \end{aligned}$$

per cui:

$$\mathbf{F}(\gamma(t)) \cdot \gamma'(t) = -\sin t \cos t + \sin t \cos t - 1 = -1.$$

Infine:

$$\int_{\gamma} \mathbf{F}(\mathbf{x}) \cdot d\mathbf{x} = - \int_0^{3\pi} dt = -3\pi.$$



La definizione di integrale di linea è invariante per archi di curva equivalenti, infatti vale:

Teorema 4.2.1 Siano γ e δ due archi regolari equivalenti e sia \mathbf{F} un campo vettoriale continuo su un aperto A contenente il loro sostegno. Allora:

$$\int_{\gamma} \mathbf{F}(\mathbf{x}) \cdot d\mathbf{x} = \int_{\delta} \mathbf{F}(\mathbf{x}) \cdot d\mathbf{x}.$$

Consideriamo, ora, un arco γ in \mathbb{R}^n regolare a tratti definito per $t \in [a, b]$. Indichiamo con $-\gamma$ l'arco:

$$(-\gamma)(u) = \gamma(-u) \quad u \in [-b, -a];$$

è evidente che se il parametro u varia da $-b$ a $-a$, il punto $(-\gamma)(u)$ sul sostegno di $-\gamma$ si muove sul sostegno di γ passando dall'estremo $\gamma(b)$ all'estremo $\gamma(a)$, ovvero il senso di percorrenza di $-\gamma$ è opposto a quello di γ .

Siano, poi, γ_1 e γ_2 due archi di curva in \mathbb{R}^n regolari a tratti tali che:

$$\gamma_1 : [a_1, b_1] \rightarrow \mathbb{R}^n \quad \gamma_2 : [a_2, b_2] \rightarrow \mathbb{R}^n$$

e caratterizzati da $\gamma_1(b_1) = \gamma_2(a_2)$. Indicheremo con $\gamma_1 + \gamma_2$ l'arco regolare a tratti ottenuto seguendo prima il percorso γ_1 e poi il percorso γ_2 .

Teorema 4.2.2 Siano γ , γ_1 e γ_2 archi di curva regolari a tratti in \mathbb{R}^n , e sia \mathbf{F} un campo vettoriale continuo su un aperto contenente i sostegni dei tre archi. Allora:

$$\begin{aligned} \int_{-\gamma} \mathbf{F}(\mathbf{x}) \cdot d\mathbf{x} &= - \int_{\gamma} \mathbf{F}(\mathbf{x}) \cdot d\mathbf{x}, \\ \int_{\gamma_1 + \gamma_2} \mathbf{F}(\mathbf{x}) \cdot d\mathbf{x} &= \int_{\gamma_1} \mathbf{F}(\mathbf{x}) \cdot d\mathbf{x} + \int_{\gamma_2} \mathbf{F}(\mathbf{x}) \cdot d\mathbf{x}. \end{aligned}$$

Tra tutti i campi vettoriali, di particolare importanza sono quelli caratterizzati dalla seguente proprietà:

Definizione 4.2.2 Un campo vettoriale $\mathbf{F}: A \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ si dice *conservativo* in un aperto A se è continuo in A ed esiste una funzione di più variabili $U : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$, detta *potenziale*, tale che $U \in C^1(A)$ e $\mathbf{F} = \nabla U$ in tutto A .

I campi conservativi godono di una importantissima proprietà:

Teorema 4.2.3 Sia \mathbf{F} un campo vettoriale continuo nell'aperto $A \subset \mathbb{R}^n$ e si supponga che $\mathbf{F}(\mathbf{x}) = \nabla U(\mathbf{x}) \forall \mathbf{x} \in A$. Allora se $\gamma: [a, b] \rightarrow A$ è un arco di curva regolare a tratti, si ha:

$$\int_{\gamma} \mathbf{F}(\mathbf{x}) \cdot d\mathbf{x} = U(\gamma(b)) - U(\gamma(a)).$$

In altri termini, l'integrale di linea di un campo conservativo non dipende dal percorso di integrazione (ovvero dall'arco γ), ma solo dai suoi estremi. Naturalmente, ciò implica che se l'arco è chiuso, l'integrale di linea di un campo conservativo è nullo.

Teorema 4.2.4 Sia $A \subset \mathbb{R}^n$ un aperto connesso, e sia \mathbf{F} un campo vettoriale continuo in A . Le seguenti proprietà sono equivalenti:

1. \mathbf{F} è conservativo;
2. se γ_1 e γ_2 sono due archi qualunque regolari a tratti con sostegno contenuto in A e aventi (nell'ordine) gli stessi estremi, allora:

$$\int_{\gamma_1} \mathbf{F}(\mathbf{x}) \cdot d\mathbf{x} = \int_{\gamma_2} \mathbf{F}(\mathbf{x}) \cdot d\mathbf{x};$$

3. se γ è un qualunque arco chiuso regolare a tratti con sostegno contenuto in A , allora:

$$\int_{\gamma} \mathbf{F}(\mathbf{x}) \cdot d\mathbf{x} = 0.$$

Teorema 4.2.5 Sia $A \subset \mathbb{R}^n$ un aperto, e sia \mathbf{F} un campo vettoriale di classe $C^1(A)$ di componenti $\mathbf{F} = (f_1, \dots, f_n)$. Condizione necessaria affinché \mathbf{F} sia conservativo è che, per ogni coppia di indici $i, j = 1, \dots, n$:

$$\frac{\partial f_j}{\partial x_i} = \frac{\partial f_i}{\partial x_j}.$$

Nel caso dei campi vettoriali in \mathbb{R}^2 , è possibile fare uso del seguente risultato, che consente di trasformare l'integrale di linea su una linea chiusa in un integrale di superficie:

Teorema 4.2.6 (Teorema di Green) Sia $\mathbf{F} = (f_1, f_2)$ un campo vettoriale di classe C^1 nell'aperto $\Omega \subset \mathbb{R}^2$. Sia A un aperto limitato contenuto in Ω con la

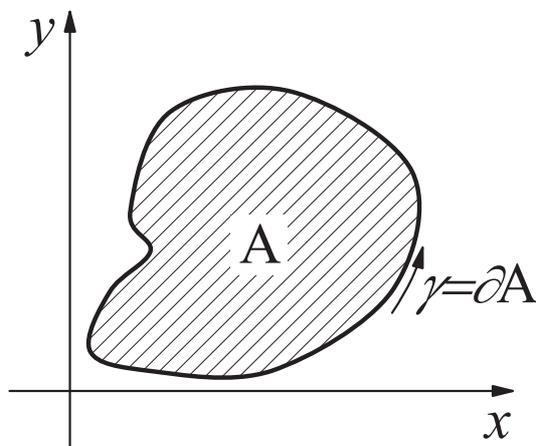


Figura 4.2. Interpretazione geometrica del senso di percorrenza antiorario per la frontiera dell'aperto A nel Teorema di Green.

sua frontiera. Si assuma che la frontiera di A sia il sostegno di un arco γ chiuso, semplice e regolare a tratti, e si supponga che γ sia percorso in senso antiorario. Allora:

$$\int_{\gamma} \mathbf{F}(\mathbf{x}) \cdot d\mathbf{x} = \int_A \left[\frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y} \right] dx dy.$$

Si noti che il Teorema di Green richiede una attenzione particolare nella scelta del senso di percorrenza della frontiera di A : il verso antiorario è tale che un osservatore, in movimento sulla curva chiusa γ , trovi l'insieme A alla sua sinistra (Fig. 4.2).

Teorema 4.2.7 (Teorema di Jordan) Sia γ un arco chiuso e semplice nel piano, di sostegno Γ . Allora $\mathbb{R}^2 \setminus \Gamma$ è l'unione di due aperti connessi disgiunti: uno di essi è limitato, e viene detto l'*interno* di γ , l'altro è illimitato e viene detto l'*esterno* di γ .

Definizione 4.2.3 Un aperto connesso $A \subset \mathbb{R}^2$ si dice *semplicemente connesso* se, dato comunque un arco chiuso γ , semplice e il cui sostegno sia contenuto in A , anche l'interno di γ è contenuto in A .

In altri termini, un aperto è semplicemente connesso solo se non contiene "buchi."

Il Teorema di Green può essere utilizzato per valutare l'area di un insieme piano delimitato da un arco chiuso, semplice e regolare a tratti per mezzo di un integrale

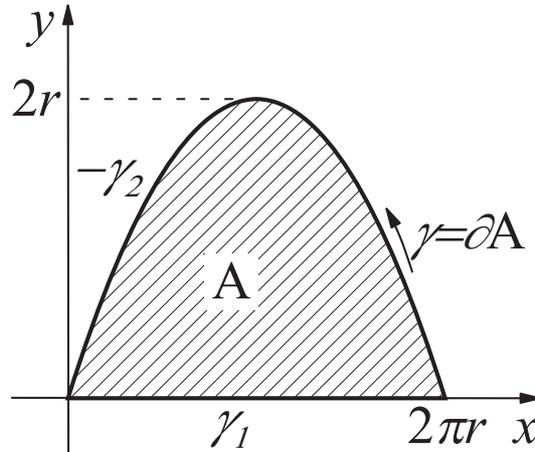


Figura 4.3. Rappresentazione grafica dell'insieme A e della sua frontiera orientata γ utilizzati nell'Esercizio 4.2.2.

di linea, purché si utilizzi un campo vettoriale \mathbf{F} di classe C^1 e tale che:

$$\frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y} = \text{cost.}$$

Esempio 4.2.2

Determinare l'area dell'insieme A compreso tra l'asse x e l'arco di curva γ_2 di equazioni:

$$\gamma_2(t) = (r(t - \sin t), r(1 - \cos t)) \quad t \in [0, 2\pi], \quad r > 0.$$

Indicando con γ_1 l'arco di curva corrispondente al tratto di asse x compreso tra l'origine e il punto $(r, 0)$:

$$\gamma_1(u) = (u, 0) \quad u \in [0, 2\pi r], \quad r > 0,$$

il percorso chiuso γ , percorso in senso antiorario, è rappresentato dall'arco regolare a tratti (si veda la Fig. 4.3):

$$\gamma = \gamma_1 + (-\gamma_2).$$

I vettori tangenti alle due curve sono quindi:

$$\gamma_1'(u) = (1, 0) \quad \gamma_2'(t) = (r(1 - \cos t), r \sin t),$$

mentre come campo vettoriale si utilizza:

$$\mathbf{F}(x, y) = (0, x) \implies \frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y} = 1.$$

Infine, per il Teorema di Green e per il Teorema 4.2.2:

$$m(A) = \int_A 1 \, dx dy = \int_{\gamma} \mathbf{F}(\mathbf{x}) \cdot d\mathbf{x} = \int_{\gamma_1} \mathbf{F}(\mathbf{x}) \cdot d\mathbf{x} - \int_{\gamma_2} \mathbf{F}(\mathbf{x}) \cdot d\mathbf{x},$$

dove:

$$\begin{aligned} \int_{\gamma_1} \mathbf{F}(\mathbf{x}) \cdot d\mathbf{x} &= \int_0^{2\pi r} (0, u) \cdot (1, 0) \, du = 0 \\ \int_{\gamma_2} \mathbf{F}(\mathbf{x}) \cdot d\mathbf{x} &= \int_0^{2\pi} (0, r(t - \sin t)) \cdot (r(1 - \cos t), r \sin t) \, dt \\ &= r^2 \int_0^{2\pi} (t - \sin t) \sin t \, dt. \end{aligned}$$

Integrando per parti si ha immediatamente:

$$\int t \sin t \, dt = -t \cos t + \int \cos t \, dt = -t \cos t + \sin t,$$

mentre per il secondo integrale, sempre integrando per parti, da:

$$\begin{aligned} \int \sin^2 t \, dt &= -\sin t \cos t + \int \cos^2 t \, dt = -\sin t \cos t + \int (1 - \sin^2 t) \, dt \\ &= -\sin t \cos t + t - \int \sin^2 t \, dt \end{aligned}$$

si ricava:

$$\int \sin^2 t \, dt = \frac{t}{2} - \frac{1}{2} \sin t \cos t.$$

Infine:

$$\int_{\gamma_2} \mathbf{F}(\mathbf{x}) \cdot d\mathbf{x} = r^2 \left[-t \cos t + \sin t + \frac{t}{2} - \frac{1}{2} \sin t \cos t \right]_0^{2\pi} = -3\pi r^2$$

e quindi si ha $m(A) = 3\pi r^2$. ◆

Il Teorema di Green, enunciato per ora solo per curve chiuse, semplici e regolari a tratti può essere esteso ad aperti A di carattere un po' più generale:

Definizione 4.2.4 Sia A un aperto connesso e limitato. Esso è un *aperto con bordo* se la sua frontiera ∂A è l'unione di un numero finito di sostegni, a due a due disgiunti, di archi chiusi, semplici e regolari a tratti $\gamma_1, \dots, \gamma_m$.

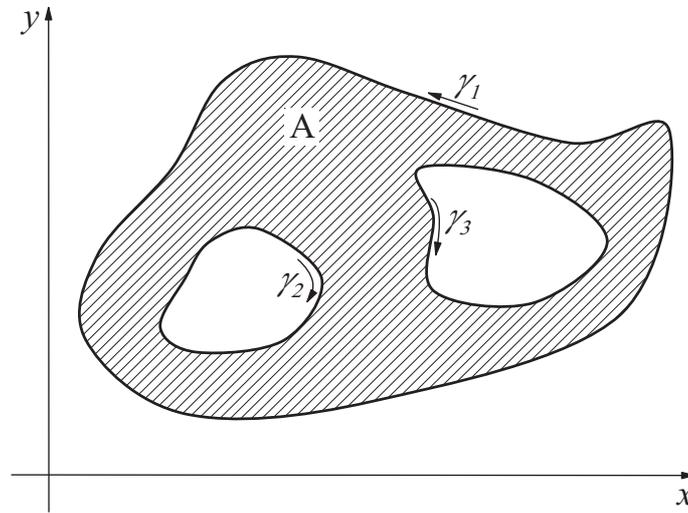


Figura 4.4. Rappresentazione grafica di un aperto con bordo A .

Si assuma che ognuno di questi archi sia orientato in modo tale che un osservatore che li percorra veda sempre l'insieme A alla sua sinistra: grazie a questa orientazione, gli archi $\gamma_1, \dots, \gamma_m$ costituiscono il *bordo* di A (che differisce dalla frontiera ∂A solo per il fatto di essere orientato, si veda la Fig. 4.4).

Teorema 4.2.8 Sia A un aperto con bordo, e siano $\gamma_1, \dots, \gamma_m$ gli archi chiusi, semplici e regolari a tratti che costituiscono il bordo di A . Sia inoltre $\mathbf{F} = (f_1, f_2)$ un campo vettoriale di classe C^1 in un aperto Ω contenente \bar{A} . Allora:

$$\sum_{i=1}^m \int_{\gamma_i} \mathbf{F}(\mathbf{x}) \cdot d\mathbf{x} = \int_A \left[\frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y} \right] dx dy.$$

Si scrive anche:

$$\int_{\partial A} \mathbf{F}(\mathbf{x}) \cdot d\mathbf{x} = \int_A \left[\frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y} \right] dx dy.$$

4.3 Integrali di flusso

Si consideri una superficie semplice e regolare $\sigma: A \rightarrow \mathbb{R}^3$, e sia γ un arco chiuso, semplice e regolare a tratti contenuto in A insieme al suo interno. Indicato con K il compatto costituito dal sostegno di γ e dal suo interno, sia σ_0 la calotta ottenuta restringendo σ a K .

Sia inoltre $\mathbf{F}(x,y,z)$ un campo vettoriale continuo sul sostegno della calotta σ_0 , e si consideri il punto $(u,v) \in K$ di immagine $\mathbf{x} = (x,y,z) = \sigma_0(u,v)$. La Definizione 4.1.15 individua il versore normale $\hat{\mathbf{n}}(u,v)$ alla calotta σ_0 nel punto \mathbf{x} .

Definizione 4.3.1 Si chiama *flusso* del campo \mathbf{F} attraverso la calotta σ_0 l'integrale su σ_0 della componente di \mathbf{F} secondo il versore normale di σ_0 :

$$\int_{\sigma_0} \mathbf{F} \cdot \hat{\mathbf{n}} = \int_K \mathbf{F}(\sigma(u,v)) \cdot \mathbf{n}(u,v) \, dudv.$$

Questa definizione è, naturalmente, indipendente dalla parametrizzazione scelta per la calotta:

Teorema 4.3.1 Siano $\sigma_0: K \rightarrow \mathbb{R}^3$ e $\rho_0: K' \rightarrow \mathbb{R}^3$ due calotte semplici equivalenti. Se \mathbf{F} è un campo vettoriale continuo sul loro comune sostegno, allora:

$$\int_{\sigma_0} \mathbf{F} \cdot \hat{\mathbf{n}} = \int_{\rho_0} \mathbf{F} \cdot \hat{\mathbf{n}}.$$

Se le due calotte σ_0 e ρ_0 , pur avendo lo stesso sostegno, ne individuano due versi di attraversamento opposti (ovvero, se esiste un cambiamento di coordinate con determinante della matrice Jacobiana negativo), si ha il seguente:

Teorema 4.3.2 Siano $\sigma_0: K \rightarrow \mathbb{R}^3$ e $\rho_0: K' \rightarrow \mathbb{R}^3$ due calotte semplici aventi lo stesso sostegno, ma con versi di attraversamento opposti, e sia \mathbf{F} un campo vettoriale continuo sul loro comune sostegno. Allora:

$$\int_{\sigma_0} \mathbf{F} \cdot \hat{\mathbf{n}} = - \int_{\rho_0} \mathbf{F} \cdot \hat{\mathbf{n}}.$$

La definizione di flusso può essere estesa in modo naturale a superfici costituite da un numero finito di calotte regolari semplici e non sovrappoventisi, purché l'orientamento di ciascuna delle calotte parziali sia compatibile con quello delle altre calotte confinanti.

Assegnato un campo vettoriale $\mathbf{F} = (f_1, f_2, f_3)$ in \mathbb{R}^3 di classe C^1 su un aperto $\Omega \subset \mathbb{R}^3$, si definisce con *rotore* di \mathbf{F} il campo vettoriale:

$$\nabla \wedge \mathbf{F} = \text{rot}\{\mathbf{F}\} = \left(\frac{\partial f_3}{\partial y} - \frac{\partial f_2}{\partial z}, \frac{\partial f_1}{\partial z} - \frac{\partial f_3}{\partial x}, \frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y} \right).$$

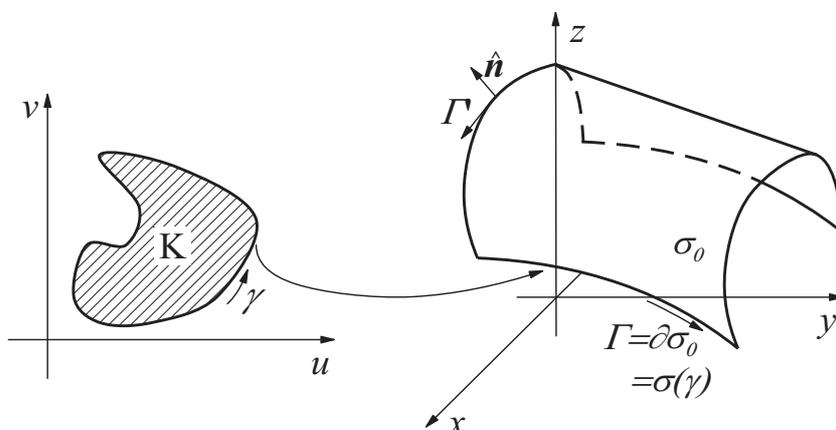


Figura 4.5. Rappresentazione grafica delle curve e delle superfici utilizzate nel Teorema di Stokes 4.3.3.

Analogamente, con *divergenza* di \mathbf{F} si intende il campo scalare:

$$\nabla \cdot \mathbf{F} = \operatorname{div}\{\mathbf{F}\} = \frac{\partial f_1}{\partial x} + \frac{\partial f_2}{\partial y} + \frac{\partial f_3}{\partial z}.$$

È possibile trasformare la circuitazione di un campo vettoriale in \mathbb{R}^3 in un flusso:

Teorema 4.3.3 (Teorema di Stokes) Siano $\Omega \subset \mathbb{R}^3$ aperto e $\mathbf{F} = (f_1, f_2, f_3) \in C^1(\Omega)$. Sia $\sigma: A \subset \mathbb{R}^2 \rightarrow \Omega$ una superficie regolare definita sull'aperto connesso A , e sia γ un arco chiuso, semplice e regolare a tratti il cui sostegno e il cui interno siano contenuti in A individuando il compatto K . Si assuma inoltre che γ sia orientata in modo da lasciare il suo interno alla sua sinistra, e sia σ_0 la restrizione di σ al compatto K . Sia, infine, $\Gamma = \sigma(\gamma) = \partial\sigma_0$ il sostegno dell'arco chiuso, semplice e regolare a tratti che costituisce il bordo³ della calotta σ_0 . Allora:

$$\int_{\partial\sigma_0} \mathbf{F}(\mathbf{x}) \cdot d\mathbf{x} = \int_{\sigma_0} \nabla \wedge \mathbf{F} \cdot \hat{\mathbf{n}}.$$

Definizione 4.3.2 Un aperto connesso $A \subset \mathbb{R}^3$ si dice *semplicemente connesso* se dato comunque un arco Γ chiuso, semplice e regolare a tratti con sostegno contenuto in A , esiste una calotta regolare σ_0 di cui Γ sia il bordo e il cui sostegno sia pure contenuto in A .

³Si noti che l'orientazione del bordo della calotta è tale che un osservatore che lo percorra vede la calotta alla sua sinistra: si veda la Fig. 4.5.

Definizione 4.3.3 Un campo vettoriale \mathbf{F} continuo nell'aperto $A \subset \mathbb{R}^3$, viene detto *irrotazionale* se $\nabla \times \mathbf{F} = \mathbf{0}$ in tutto A .

Teorema 4.3.4 Sia $A \subset \mathbb{R}^3$ un aperto semplicemente connesso e sia \mathbf{F} un campo vettoriale di classe $C^1(A)$ e irrotazionale. Allora \mathbf{F} è conservativo.

Sia, ora, \mathbf{F} un campo vettoriale definito in un aperto $\Omega \subset \mathbb{R}^3$, e sia $A \subset \Omega$ un aperto limitato, la cui frontiera ∂A sia l'unione di un numero finito di sostegni di calotte regolari e semplici $\sigma_1, \dots, \sigma_n$, intersecantisi solo per punti dei loro bordi. Se ciascuna di tali calotte viene orientata in modo che il relativo versore normale punti verso l'esterno di A , si dice che tali calotte costituiscono il *bordo* di A , e si dice anche che A è un *aperto con bordo*. Indichiamo il *flusso totale uscente* da A con:

$$\int_{\partial A} \mathbf{F} \cdot \hat{\mathbf{n}} = \sum_{i=1}^n \int_{\sigma_i} \mathbf{F} \cdot \hat{\mathbf{n}}.$$

Se il campo vettoriale è sufficientemente regolare, questo flusso può essere calcolato come un integrale di volume:

Teorema 4.3.5 (Teorema di Gauss) Sia \mathbf{F} un campo vettoriale di classe C^1 nell'aperto $\Omega \subset \mathbb{R}^3$, e sia $A \subset \Omega$ un aperto limitato con bordo. Allora:

$$\int_{\partial A} \mathbf{F} \cdot \hat{\mathbf{n}} = \int_A \nabla \cdot \mathbf{F} \, dx dy dz.$$

Esempio 4.3.1

Calcolare il flusso totale uscente dalla superficie sferica di centro l'origine e raggio $r > 0$ per il campo $\mathbf{F} = (x, 0, y)$.

Dalla definizione di divergenza si ha:

$$\nabla \cdot \mathbf{F} = \frac{\partial f_1}{\partial x} + \frac{\partial f_2}{\partial y} + \frac{\partial f_3}{\partial z} = 1,$$

pertanto, per il Teorema di Gauss:

$$\int_{\partial A} \mathbf{F} \cdot \hat{\mathbf{n}} = \int_A 1 \, dx dy dz = \frac{4}{3} \pi r^3,$$

essendo A la sfera di centro l'origine e raggio r . ◆

5

Calcolo delle probabilità

Obiettivi: Elementi di teoria della probabilità. Calcolo combinatorio. Variabili casuali discrete. Variabili casuali continue.

5.1 Elementi di teoria della probabilità

La teoria della probabilità, nato nel XVI secolo a seguito di un tentativo di dare un fondamento matematico alla teoria dei giochi, è poi stati sistematizzato nella attuale forma assiomatica solo nel 1933 da parte del matematico russo Kolmogorov. Il primo concetto fondamentale della teoria della probabilità è quello di *esperimento casuale*: si tratta di un esperimento suscettibile di più risultati diversi, ma il cui risultato non possa essere predetto con certezza.

Si consideri, quindi, un esperimento casuale, e si assuma di poterlo effettuare un numero arbitrario di volte nelle stesse condizioni. Oggetto della teoria della probabilità è l'individuazione di un insieme di strumenti matematici che consentano di descrivere accuratamente i risultati di un gran numero di tali eventi. L'esperimento casuale viene descritto facendo uso della teoria degli insiemi; in particolare, si costruisce un insieme, detto *spazio campione* Ω , i cui elementi sono tutti e soli i possibili risultati dell'esperimento casuale in esame.

Esempio 5.1.1

Si consideri il lancio di una moneta, che presenta i due risultati possibili T (testa) e C (croce). Allora lo spazio campione sarà:

$$\Omega = \{T, C\}.$$

Poiché la cardinalità di Ω è finita, si parla di *spazio campione finito*. ♦

Esempio 5.1.2

Un altro esempio di spazio campione finito è dato dall'esperimento casuale corrispondente al lancio di due monete, cui corrispondono 4 possibili risultati:

$$\Omega = \{TT, CC, TC, CT\}.$$

♦

Esempio 5.1.3

Si consideri nuovamente il lancio di una moneta, che però ora viene ripetuto fino a che non si ottiene il risultato T . Allora, lo spazio campione è costituito da una infinità numerabile di risultati possibili:

$$\Omega = \{s_i\}_{i=1}^{\infty},$$

dove il generico elemento s_i corrisponde a:

$$s_i = \underbrace{CC \cdots C}_{i-1} T.$$

Poiché Ω ha la cardinalità di \mathbb{N} , si parla di *spazio campione infinito numerabile*. ♦

Esempio 5.1.4

Si consideri l'esperimento casuale corrispondente all'arrivo di un treno in un istante di tempo compreso tra 0 e 60 minuti. Allora lo spazio campione è l'intervallo $\Omega = I = [0,60] \subset \mathbb{R}$, e quindi ha la cardinalità di \mathbb{R} . Si parla di *spazio campione continuo*. ♦

Una volta costruito lo spazio campione, si considerano tutti i suoi sottoinsiemi: ogni sottoinsieme dello spazio campione costituisce un *evento*. Pertanto, un evento è ogni risultato, comunque complesso, che si può presentare a conclusione dell'esperimento casuale. In termini dello spazio campione, un evento è rappresentato dalla collezione dei punti di Ω associati a quei risultati dell'esperimento per cui l'evento stesso è verificato. Dato un evento $A \subset \Omega$, anche $\Omega \setminus A$ è un evento, corrispondente al non verificarsi dell'evento A stesso.

Esempio 5.1.5

Si consideri l'esperimento casuale costituito dal lancio di un dado, il cui spazio campione finito è costituito dai sei numeri che corrispondono alle sei facce del dado:

$$\Omega = \{1,2,3,4,5,6\}.$$

L'evento "uscita di un numero dispari" corrisponde al sottoinsieme:

$$A = \{1,3,5\} \subset \Omega.$$



Tra tutti i sottoinsiemi di Ω , si hanno:

1. lo stesso spazio campione Ω : poiché questo è stato costruito inserendovi tutti i possibili risultati dell'esperimento casuale, esso viene anche detto *evento certo*;
2. l'insieme vuoto \emptyset : poiché questo non contiene alcun punto di Ω , non si verifica mai, e viene quindi detto *evento impossibile*.

Dati due eventi $A, B \subset \Omega$, si definiscono gli eventi:

1. $A \cup B$, corrispondente al verificarsi di A , di B , oppure di entrambi;
2. $A \cap B$, corrispondente al verificarsi contemporaneo di A e di B .

Nel caso in cui $A \cap B = \emptyset$, si dice che i due eventi A e B sono *disgiunti*, o *mutuamente esclusivi*.

Una volta costruito lo spazio delle probabilità, si associa ad ogni evento $A \subset \Omega$ un numero $P(A)$, compreso tra 0 e 1, chiamato *probabilità* dell'evento A . In altri termini, si definisce la *funzione*:

$$P: \Omega \rightarrow [0,1] \subset \mathbb{R}.$$

La definizione di $P(A)$ deve soddisfare un certo numero di assiomi, scelti in modo tale da adeguarsi sia all'intuito sia all'evidenza sperimentale. In particolare, la probabilità viene spesso interpretata come caso limite della *frequenza statistica*: questa corrisponde al rapporto tra il numero di risultati favorevoli $n(A)$ dell'esperimento casuale e il numero totale N di esperimenti effettuati, dove naturalmente l'esperimento è favorevole qualora si verifichi l'evento A considerato. Gli assiomi che la probabilità deve verificare sono:

1. $P(A) \geq 0 \forall A \subset \Omega$;
2. $P(\Omega) = 1$;
3. se $A, B \subset \Omega$ sono disgiunti, allora $P(A \cup B) = P(A) + P(B)$.

È immediato verificare come la frequenza statistica soddisfi, qualunque sia N finito, questi assiomi.

Resta ancora da risolvere il problema di come assegnare ad ogni evento A la sua probabilità. Si può pensare di condurre delle misure ripetute di frequenza statistica effettuando un grande numero di prove, oppure si può tentare di costruire lo spazio campione Ω in modo tale che tutti i suoi elementi consentano una assegnazione di probabilità in modo intuitivo. In particolare, nel caso di uno spazio campione finito di cardinalità N è molto comodo costruire Ω in modo tale che ad ogni elemento ω sia associata la stessa probabilità¹ p . In questo caso, si ha evidentemente:

$$1 = P(\Omega) = \sum_{\omega \in \Omega} P(\{\omega\}) = Np,$$

da cui si ricava immediatamente:

$$p = \frac{1}{N}.$$

Se, poi, si considera un generico evento $A \subset \Omega$ costituito da $n(A) \in [0, N] \subset \mathbb{N}$ elementi dello spazio campione, si dice che $n(A)$ è il *numero di eventi favorevoli* di A . Per le precedenti considerazioni, si ha:

$$P(A) = n(A)p = \frac{n(A)}{N}. \quad (5.1)$$

Quindi la valutazione della probabilità $P(A)$ richiede di essere in grado di contare il numero di eventi favorevoli dell'evento in questione per poi rapportarlo al numero totale N di eventi possibili, in accordo con la (5.1). Quando $n(A)$ è molto grande, la costruzione esplicita dello spazio campione e il conteggio degli eventi favorevoli possono risultare impossibili: si fa allora uso del calcolo combinatorio, presentato nella Sez. 5.2.

Esempio 5.1.6

Si consideri nuovamente l'esperimento casuale costituito dal lancio di un dado, già discusso nell'Esempio 5.1.5. Lo spazio campione ha cardinalità $N = 6$. Nell'ipotesi di non avere a che fare con un dado truccato, si può ritenere che ognuno dei sei eventi sia equiprobabile, e quindi $p = 1/6$. La probabilità associata all'evento "uscita di un numero dispari," corrispondente al sottoinsieme:

$$A = \{1, 3, 5\} \subset \Omega$$

¹Per essere più precisi, dato un generico elemento $\omega \in \Omega$, si intende che la probabilità associata all'evento $\{\omega\} \subset \Omega$ sia indipendente dall'elemento ω stesso.

di cardinalità $n(A) = 3$, vale quindi $P(A) = n(A)/N = 1/2$. \blacklozenge

A partire dagli assiomi della probabilità, si dimostra facilmente:

Teorema 5.1.1 Sia Ω uno spazio campione dotato della funzione di probabilità P , e siano $A, B \subset \Omega$ due eventi di tale spazio campione. Allora si ha:

1. $P(\emptyset) = 0$;
2. $P(\Omega \setminus A) = 1 - P(A)$;
3. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Consideriamo, ora, due eventi $A, B \subset \Omega$ nello spazio campione Ω . L'evento $A \cap B$ corrisponde al verificarsi congiuntamente dei due eventi parziali A e B : la probabilità ad esso associata viene detta *probabilità congiunta* $P(A, B)$ dei due eventi:

$$P(A, B) = P(A \cap B).$$

Poiché l'intersezione di due insiemi è una operazione commutativa, è immediato verificare che $P(A, B) = P(B, A)$.

Esempio 5.1.7

Si consideri il lancio indipendente di due dadi, e si indichi con s_{ij} un elemento dello spazio campione costituito dall'uscita del numero i per il primo dado, e j per il secondo. Evidentemente, lo spazio campione sarà costituito da 36 elementi. Si considerino, ora, i due eventi:

$$A = \{s_{11}, s_{12}, s_{13}, s_{14}, s_{15}, s_{16}\}, \quad B = \{s_{13}, s_{23}, s_{33}, s_{43}, s_{53}, s_{63}\},$$

che corrispondono, rispettivamente, ad ottenere 1 dal primo dado e 3 dal secondo. La loro intersezione, quindi, corrisponde ad ottenere 1 dal primo dado e 3 dal secondo:

$$A \cap B = \{s_{13}\},$$

pertanto, assumendo che ogni elemento dello spazio campione sia equiprobabile, la probabilità congiunta vale $P(A, B) = 1/36$. Si noti, inoltre, che $P(A) = P(B) = 6/36 = 1/6$, e quindi $P(A, B) = P(A)P(B)$. \blacklozenge

Definizione 5.1.1 Si consideri uno spazio campione Ω , cui sia associata una certa funzione di probabilità P . Gli eventi $A, B \subset \Omega$ si dicono *statisticamente indipendenti* se:

$$P(A, B) = P(A)P(B).$$

Definizione 5.1.2 Si consideri uno spazio campione Ω , cui sia associata una certa funzione di probabilità P . Gli n eventi $A_1, A_2, \dots, A_n \subset \Omega$ si dicono *a due a due statisticamente indipendenti* se per ogni scelta di indici $i, j = 1, \dots, n$ con $i \neq j$ si ha:

$$P(A_i, A_j) = P(A_i)P(A_j).$$

Si dice, invece, che formano una *famiglia di eventi statisticamente indipendenti* se per ogni $k \leq n$ e per ogni scelta di indici i_1, \dots, i_k tutti distinti e compresi tra 1 e n si ha:

$$P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}) = P(A_{i_1}, A_{i_2}, \dots, A_{i_k}) = \prod_{j=1}^k P(A_{i_j}).$$

Si noti come sia impossibile, dalla sola conoscenza dello spazio campione, decidere in generale dell'indipendenza statistica di due eventi. Per poterla affermare occorre fare uso di considerazioni fisiche sull'esperimento casuale in oggetto, oppure dei risultati del calcolo. Inoltre possono esistere eventi a due a due indipendenti ma che non formano una famiglia di eventi indipendenti:

Esempio 5.1.8

Si consideri lo spazio campione costituito da elementi equiprobabili $\Omega = \{1, 2, 3, 4\}$, e si considerino i tre eventi:

$$A_1 = \{1, 4\}, \quad A_2 = \{2, 4\}, \quad A_3 = \{3, 4\}.$$

Essi sono indipendenti a due a due, infatti:

$$P(A_1, A_2) = P(A_1, A_3) = P(A_2, A_3) = P(\{4\}) = \frac{1}{4}$$

e $P(A_1) = P(A_2) = P(A_3) = 1/2$, ma non formano una famiglia di eventi statisticamente indipendenti, infatti:

$$P(A_1, A_2, A_3) = P(\{4\}) = \frac{1}{4} \neq \frac{1}{8} = P(A_1)P(A_2)P(A_3).$$



Esempio 5.1.9 (Schema di Bernoulli)

Un tipo molto comune di esperimento casuale consiste nel ripetere un numero n di volte, in modo indipendente,² un certo esperimento, i cui risultati possano essere solo due. Convenzionalmente, i due risultati si indicano con 1 (successo) e 0 (fallimento). Si indichi con p la probabilità che ogni singolo esperimento dia risultato 1, e quindi $1 - p$ sia la probabilità che il risultato sia 0. Si noti che questa scelta non corrisponde ad uno spazio di probabilità con elementi equiprobabili, a meno che $p = 1/2$. Vogliamo determinare la probabilità di avere una certa sequenza prefissata di n elementi costituita da k 1 e $(n - k)$ 0.

Faremo uso dello spazio di probabilità costituito da tutte le possibili n -uple di elementi 0 e 1:

$$\Omega = \{\omega = (\omega_1, \dots, \omega_n) \text{ con } \omega_i = 0 \text{ oppure } \omega_i = 1 \text{ per } i = 1, \dots, n\}.$$

La sequenza di cui vogliamo determinare la probabilità sarà:

$$\omega_0 = (\underbrace{1, \dots, 1}_{k \text{ volte}}, \underbrace{0, \dots, 0}_{n - k \text{ volte}}).$$

Sia $A_i = \{\omega \text{ con } \omega_i = 1\}$, per $i = 1, \dots, n$, il sottoinsieme di Ω corrispondente all'evento "il risultato dell' i -esimo esperimento ha risultato 1", cui evidentemente corrisponde $P(A_i) = p$. Allora, l'evento $\{\omega_0\}$ potrà essere scritto:

$$\{\omega_0\} = A_1 \cap \dots \cap A_k \cap (\Omega \setminus A_{k+1}) \cap \dots \cap (\Omega \setminus A_n).$$

Poiché non c'è motivo di pensare che la conoscenza dei risultati di alcune ripetizioni dell'esperimento influenzi il risultato delle altre, possiamo assumere che gli n eventi A_i formino una famiglia di eventi statisticamente indipendenti, e quindi:

$$P(\{\omega_0\}) = P(A_1) \dots P(A_k) [1 - P(A_{k+1})] \dots [1 - P(A_n)] = p^k (1 - p)^{n-k}.$$

È immediato convincersi che questo risultato corrisponde alla probabilità di avere k risultati 1 e $n - k$ risultati 0 in n esperimenti, indipendentemente dalla posizione dei risultati stessi nella sequenza di n risultati. \blacklozenge

Di grande importanza pratica è anche il caso in cui si voglia stimare la probabilità del verificarsi di un evento A condizionata dall'occorrenza preventiva di un altro evento B . Si parla, in questo caso, di *probabilità dell'evento A condizionata all'evento B* , indicata con $P(A|B)$.

²Ad esempio, l'indipendenza si può ottenere ripristinando ogni volta le stesse condizioni iniziali dell'esperimento.

Definizione 5.1.3 Sia Ω uno spazio campione dotato di una funzione di probabilità P , e siano $A, B \subset \Omega$ due eventi di Ω con $P(B) \neq 0$. Si definisce la *probabilità condizionata* $P(A|B)$ come:

$$P(A|B) = \frac{P(A, B)}{P(B)}.$$

Si noti che, se i due eventi A e B sono statisticamente indipendenti, si ha immediatamente:

$$P(A|B) = \frac{P(A, B)}{P(B)} = \frac{P(A)P(B)}{P(B)} = P(A).$$

Teorema 5.1.2 Sia Ω uno spazio campione dotato di una funzione di probabilità P , e siano $A, B \subset \Omega$ due eventi di Ω con $P(B) \neq 0$. Allora A e B sono statisticamente indipendenti se e solo se $P(A|B) = P(A)$.

Teorema 5.1.3 (Teorema della probabilità totale) Sia Ω uno spazio campione dotato di una funzione di probabilità P , e siano $A_1, \dots, A_n \subset \Omega$ n eventi a due a due mutuamente esclusivi ed esaustivi, ovvero tali che:

$$A_i \cap A_j = \emptyset \text{ per } i \neq j, \quad \bigcup_{i=1}^n A_i = \Omega.$$

Allora, la probabilità di un qualunque evento $B \subset \Omega$ si può esprimere secondo:

$$P(B) = \sum_{i=1}^n P(B, A_i) = \sum_{i=1}^n P(B|A_i)P(A_i).$$

Teorema 5.1.4 (Teorema di Bayes) Sia Ω uno spazio campione dotato di una funzione di probabilità P , e siano $A_1, \dots, A_n \subset \Omega$ n eventi a due a due mutuamente esclusivi ed esaustivi. Allora, per un qualunque evento $B \subset \Omega$ e per ogni $i = 1, \dots, n$, si ha:

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{P(B)} = \frac{P(B|A_i)P(A_i)}{\sum_{i=1}^n P(B|A_i)P(A_i)}.$$

5.2 Calcolo combinatorio

Il calcolo combinatorio serve a determinare la cardinalità degli insiemi finiti, e quindi è particolarmente utile quando si calcolano probabilità di eventi in spazi campione costituiti da elementi equiprobabili.

Si considerino due insiemi N, K di cardinalità, rispettivamente, n e k .

Teorema 5.2.1 L'insieme prodotto cartesiano $N \times K$ ha cardinalità nk . Invece, l'insieme

$$N^m = \underbrace{N \times \dots \times N}_{m \text{ volte}}$$

ha cardinalità n^m .

Teorema 5.2.2 Sia $k \leq n$. Sia D_n^k l'insieme delle applicazioni iniettive $f: K \rightarrow N$. Allora D_n^k ha cardinalità $n!/(n-k)!$.

Si noti che una applicazione iniettiva da K in N corrisponde a scegliere una k -upla ordinata di elementi di N tutti distinti tra loro, ovvero a scegliere una *disposizione* di k elementi di N . Pertanto, la cardinalità di D_n^k corrisponde al **numero di possibili disposizioni degli n elementi di N in sequenze di k elementi distinti tra loro**.

Consideriamo, ora, $N = K$. Una applicazione iniettiva di N in se stesso viene detta *permutazione* di elementi di N . Per il Teorema 5.2.2, **l'insieme delle permutazioni di N ha cardinalità $n!$** .

Teorema 5.2.3 Sia C_n^k l'insieme dei sottoinsiemi di N di cardinalità $k \leq n$. Allora la cardinalità di C_n^k è data dal *binomiale*:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Quest'ultimo risultato consente di contare il **numero di insiemi non ordinati** (quelli ordinati sono contati sulla base del Teorema 5.2.2) **di k elementi scelti tra n elementi**. In altri termini, si utilizza il binomiale quando è necessario contare il numero di campioni che siano indistinguibili per permutazioni del solo ordine dei loro elementi. In questa interpretazione, il Teorema 5.2.3 può essere giustificato nel modo seguente. Il numero di k -uple ordinate (quindi, due elementi che differiscano per una permutazione di due elementi sono distinguibili) è dato dal Teorema 5.2.2: $n!/(n-k)!$. All'interno di questo numero, possiamo suddividere i campioni in classi, ciascuna delle quali è formata prendendo i campioni che contengono, in ordine diverso, gli stessi k elementi. Quindi, ciascuna classe contiene tanti elementi quante sono le permutazioni di k elementi, cioè $k!$. Il numero delle classi, infine, sarà dato dal

rapporto:

$$\frac{n!/(n-k)!}{k!} = \binom{n}{k}.$$

Esempio 5.2.1 (Formula di Bernoulli)

Si consideri nuovamente l'esperimento casuale dell'Esempio 5.1.9, nel quale i due possibili risultati, indicati convenzionalmente con 1 e 0, hanno, rispettivamente, probabilità p e $1-p$ di presentarsi. Nell'Esempio 5.1.9 abbiamo valutato la probabilità che, effettuando n esperimenti, i primi k abbiano risultato 1 e gli ultimi $n-k$ abbiano risultato 0:

$$P(\{\omega_0\}) = p^k(1-p)^{n-k}.$$

Vogliamo, ora, determinare la probabilità $P_{k,n}$ che, degli n esperimenti, k abbiano risultato 1, indipendentemente dall'ordinamento. Dobbiamo, quindi, contare il numero di modi possibili in cui la sequenza di k successi si può presentare sulle n prove. In altri termini, dobbiamo determinare il numero di eventi dello spazio campione che differiscano da ω_0 solo per una permutazione degli stessi elementi. Tale numero è dato dal Teorema 5.2.3, e vale:

$$\binom{n}{k}.$$

Tenendo conto che $P(\{\omega_0\}) = p^k(1-p)^{n-k}$, si ha immediatamente:

$$P_{k,n} = \binom{n}{k} p^k (1-p)^{n-k}.$$



Esempio 5.2.2 (Legge ipergeometrica)

Calcolare la probabilità di estrarre contemporaneamente, da un sacchetto contenente $n = 7$ biglie delle quali $n_1 = 4$ sono rosse e $n - n_1 = 3$ sono blu, una biglia rossa e una blu.

Il numero di modi equivalenti di estrarre $r = 2$ biglie da n è:

$$\binom{n}{r} = \binom{7}{2},$$

che costituisce la cardinalità dello spazio campione, e quindi anche il denominatore della probabilità richiesta. Al numeratore, occorre fare uso del numero di eventi favorevoli. Questo è dato dal prodotto del numero di eventi corrispondenti all'estrazione

di $k = 1$ biglie rosse delle n_1 disponibili nel sacchetto:

$$\binom{n_1}{k} = \binom{4}{1},$$

e all'estrazione di $r - k = 1$ biglie blu delle $n - n_1$ disponibili nel sacchetto:

$$\binom{n - n_1}{r - k} = \binom{3}{1}.$$

In definitiva, la probabilità richiesta vale quindi (formula *ipergeometrica*):

$$P = \frac{\binom{n_1}{k} \binom{n - n_1}{r - k}}{\binom{n}{r}} = \frac{\binom{4}{1} \binom{3}{1}}{\binom{7}{2}}.$$



Esempio 5.2.3

Calcolare la probabilità di estrarre 2 assi da un mazzo di 52 carte se si scelgono 6 carte.

Si tratta di una applicazione della legge ipergeometrica, infatti il numero di modi possibili di estrarre 6 carte da un mazzo di 52 è:

$$\binom{52}{6}.$$

Nel mazzo sono presenti 4 assi, per cui il numero di modi possibili di estrarne 2 vale:

$$\binom{4}{2}.$$

Infine, i modi possibili di estrarre le altre 4 carte dalle restanti 48 del mazzo sono:

$$\binom{48}{4}.$$

In definitiva, la probabilità richiesta vale:

$$\frac{\binom{4}{2} \binom{48}{4}}{\binom{52}{6}}.$$



Esempio 5.2.4

Si calcoli la probabilità di estrarre, da un mazzo di 52, 13 carte delle quali 5 siano cuori, 4 fiori, 3 picche e 1 quadri. Al solito, il denominatore della probabilità richiesta corrisponde al numero di modi di estrarre 13 carte da 52:

$$\binom{52}{13}.$$

Per quanto riguarda i semi, le 52 carte sono ripartite in 13 carte per ogni seme. Allora il numero di modi equivalenti di estrarre 5 cuori vale:

$$\binom{13}{5}$$

mentre per gli altri semi si ha, rispettivamente:

$$\binom{13}{4}, \quad \binom{13}{3}, \quad \binom{13}{1}.$$

Infine:

$$P = \frac{\binom{13}{5} \binom{13}{4} \binom{13}{3} \binom{13}{1}}{\binom{52}{13}}.$$

◆

Esempio 5.2.5

In un'urna sono contenute 8 palline rosse, 3 bianche e 12 verdi. Si estraggono contemporaneamente 2 palline. Calcolare la probabilità che siano di colore diverso.

Per calcolare la probabilità $P(\text{diverse})$ conviene valutare la probabilità complementare corrispondente all'evento "estrazione di due palline dello stesso colore," indicata con $P(\text{uguali})$. L'evento richiesto ha luogo se si estraggono due palline rosse, o due bianche, o due verdi, eventi tra loro mutuamente esclusivi:

$$P(\text{uguali}) = P(\text{entrambe rosse}) + P(\text{entrambe bianche}) + P(\text{entrambe verdi}).$$

Date 23 palline, l'estrazione di 2 può avvenire in

$$\binom{23}{2}$$

modi diversi, mentre i modi possibili di estrarre 2 palline rosse su 8 sono:

$$\binom{8}{2},$$

e analogamente per le 3 palline bianche e le 12 verdi si ha:

$$\binom{3}{2}, \quad \binom{12}{2}.$$

Infine, si ha:

$$P(\text{uguali}) = \frac{\binom{8}{2}}{\binom{23}{2}} + \frac{\binom{3}{2}}{\binom{23}{2}} + \frac{\binom{12}{2}}{\binom{23}{2}},$$

e quindi:

$$P(\text{diverse}) = 1 - P(\text{uguali}) = 1 - \frac{\binom{8}{2} + \binom{3}{2} + \binom{12}{2}}{\binom{23}{2}}.$$



Esempio 5.2.6

Si consideri un ufficio frequentato da 100 persone, di cui 60 sono uomini e 40 donne. È noto che il 30% degli uomini e il 65% delle donne fumano. Quale è la probabilità che, dei 100 colleghi di lavoro, un fumatore sia un uomo?

Lo spazio degli eventi può essere rappresentato dai primi 100 numeri interi:

$$\Omega = \{1, 2, 3, 4, \dots, 99, 100\},$$

per cui gli insiemi degli uomini U e delle donne D possono essere identificati con i sottoinsiemi:

$$U = \{1, \dots, 60\} \subset \Omega, \quad D = \{61, \dots, 100\} \subset \Omega.$$

Indicando con F l'insieme dei fumatori, le ipotesi dell'Esercizio corrispondono a:

$$P(U) = 0.6, \quad P(D) = 0.4, \quad P(F|U) = 0.3, \quad P(F|D) = 0.65.$$

La probabilità richiesta è $P(U|F)$. Per il Teorema di Bayes 5.1.4, si ha:

$$P(U|F) = \frac{P(F|U)P(U)}{P(F)},$$

dove la probabilità associata all'essere un fumatore può venire valutata sulla base del Teorema della probabilità totale 5.1.3, infatti gli eventi U e D sono disgiunti ed esaustivi:

$$P(F) = P(F|U)P(U) + P(F|D)P(D) = 0.3 \cdot 0.6 + 0.65 \cdot 0.4 = 0.44.$$

Infine, si ha:

$$P(U|F) = \frac{0.3 \cdot 0.6}{0.44} = 0.409.$$



5.3 Variabili casuali discrete

Una volta introdotti i concetti di esperimento casuale, spazio campione e probabilità, è opportuno associare ad ogni possibile risultato dell'esperimento casuale un numero reale, ovvero definire una funzione sullo spazio campione Ω :

Definizione 5.3.1 Sia Ω uno spazio campione dotato della funzione di probabilità P . Si chiama *variabile casuale* o *variabile aleatoria* una applicazione $\xi: \Omega \rightarrow \mathbb{R}$ tale che, per ogni $x \in \mathbb{R}$, l'insieme $\{\omega \in \Omega : \xi(\omega) \leq x\}$ sia un evento di Ω .

In altri termini, la condizione fondamentale per poter definire correttamente una variabile aleatoria è che sia possibile valutare la probabilità che essa assuma valori non superiori a x , qualunque sia $x \in \mathbb{R}$. Più in generale, quindi, risulta fondamentale il calcolo di probabilità del tipo $P(\{\omega \in \Omega : \xi(\omega) \in A \subset \mathbb{R}\})$, ovvero studiare l'applicazione:

$$A \subset \mathbb{R} \rightarrow P(\{\omega \in \Omega : \xi(\omega) \in A\}) \in [0,1],$$

che ad ogni sottoinsieme di \mathbb{R} associa la probabilità che la variabile casuale ξ prenda valori in A . Questa applicazione viene detta *legge* o *distribuzione* di ξ .

In generale, non si può definire la distribuzione di una variabile casuale ξ qualunque sia il sottoinsieme $A \subset \mathbb{R}$, nel senso che potrebbe accadere che $\{\omega \in \Omega : \xi(\omega) \in A\}$ non sia un evento in Ω . In pratica, però, la classe di sottoinsiemi di \mathbb{R} per cui la legge di ξ risulta ben definita è abbastanza vasta. In particolare, se ξ è una variabile aleatoria, sono eventi in Ω :

- $\{\omega \in \Omega : \xi(\omega) > x_0\} \forall x_0 \in \mathbb{R}$, in quanto insieme complementare di $\{\omega \in \Omega : \xi(\omega) \leq x_0\}$;

- $\{\omega \in \Omega : x_0 < \xi(\omega) \leq x_1\} \forall x_0, x_1 \in \mathbb{R}, x_0 < x_1$, in quanto intersezione dei due eventi:

$$\{\omega \in \Omega : x_0 < \xi(\omega) \leq x_1\} = \{\omega \in \Omega : \xi(\omega) > x_0\} \cap \{\omega \in \Omega : \xi(\omega) \leq x_1\};$$

- $\{\omega \in \Omega : \xi(\omega) = x_0\} \forall x_0 \in \mathbb{R}$, poiché è possibile ottenerlo come intersezione di eventi mediante la relazione:

$$\{\omega \in \Omega : \xi(\omega) = x_0\} = \bigcap_n \left\{ \omega \in \Omega : x_0 - \frac{1}{n} < \xi(\omega) \leq x_0 \right\}.$$

Per brevità, utilizzeremo nel seguito la notazione:

$$\{\xi = x\} = \{\omega \in \Omega : \xi(\omega) = x\},$$

e analogamente per le altre relazioni d'ordine, ad esempio:

$$\{\xi \leq x\} = \{\omega \in \Omega : \xi(\omega) \leq x\}.$$

Sia ξ una variabile casuale, e sia inoltre $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ una funzione. Si può dimostrare che, se φ è sufficientemente regolare, $\eta = \varphi(\xi)$ è anch'essa una variabile casuale: si parla, in questo caso, di una *funzione di variabile casuale*. La definizione può essere generalizzata alle *funzioni di più variabili casuali* considerando delle funzioni φ di più variabili. Ad esempio, se ξ_1 e ξ_2 sono variabili casuali, anche $\eta = \varphi(\xi_1, \xi_2) = \xi_1 + \xi_2$ è una variabile aleatoria.

Le variabili casuali possono essere classificate in due famiglie:

- variabili casuali *discrete*, se possono assumere valori in un sottoinsieme numerabile di \mathbb{R} ;
- variabili aleatorie *continue*, se possono assumere valori in un sottoinsieme di \mathbb{R} con la cardinalità di \mathbb{R} .

Nel resto di questa Sezione, prenderemo in considerazione solamente variabili casuali discrete. In particolare, se ξ è una variabile aleatoria discreta indicheremo con $\{x_1, x_2, \dots, x_n, \dots\}$ l'insieme, al più numerabile, di valori che essa può assumere³.

³Evidentemente, si avrà $x_i \neq x_j$ per ogni $i \neq j$.

Definizione 5.3.2 Siano ξ_1, \dots, ξ_n n variabili casuali discrete definite sullo stesso spazio campione Ω dotato della funzione di probabilità P . Esse vengono dette *indipendenti* se per qualunque scelta di $A_1, \dots, A_m \subset \mathbb{R}$, $m \leq n$, si ha:

$$P(\{\xi_1 \in A_1, \dots, \xi_m \in A_m\}) = P(\{\xi_1 \in A_1\}) \dots P(\{\xi_m \in A_m\}).$$

Se le variabili aleatorie discrete sono in un numero infinito, la definizione di indipendenza si estende richiedendo l'indipendenza di qualunque loro sottoinsieme finito.

Teorema 5.3.1 Siano ξ_1, \dots, ξ_n n variabili casuali discrete indipendenti, e sia φ una funzione di variabile casuale. Allora anche le variabili casuali $\eta_1 = \varphi(\xi_1), \dots, \eta_n = \varphi(\xi_n)$ sono indipendenti.

5.3.1 Densità discreta e funzione di ripartizione

Definizione 5.3.3 Data una variabile casuale discreta ξ , diremo *densità discreta* di ξ la funzione $p_\xi(x) = P(\{\omega \in \Omega : \xi(\omega) = x\})$.

La densità discreta gode delle seguenti proprietà:

1. $p(x) = 0$ in \mathbb{R} , tranne al più una infinità numerabile di valori di $x \in \mathbb{R}$, coincidenti con i valori presi da ξ ;
2. poiché gli eventi $\omega_i = \{\omega \in \Omega : \xi(\omega) = x_i\}$ sono disgiunti ed esaustivi, si ha:

$$\sum_i p_\xi(x_i) = \sum_i P(\omega_i) = P\left(\bigcup_i \omega_i\right) = P(\Omega) = 1,$$

e quindi la densità discreta gode della proprietà:

$$\sum_{x \in \mathbb{R}} p_\xi(x) = 1,$$

infatti $p_\xi(x)$ è nulla per tutti gli $x \notin \{x_1, x_2, \dots, x_n, \dots\}$.

La conoscenza della densità discreta consente di determinare facilmente la distribuzione di ξ , infatti dato $A \subset \mathbb{R}$ l'evento:

$$\omega_A = \{\omega \in \Omega : \xi(\omega) \in A \cap \{x_1, x_2, \dots, x_n, \dots\}\}$$

può essere scritto come unione degli eventi disgiunti:

$$\omega_A = \bigcup_{x_i \in A} \{\xi = x_i\},$$

e quindi:

$$P(\omega_A) = \sum_{x_i \in A} P(\{\xi = x_i\}) = \sum_{x_i \in A} p_\xi(x_i). \quad (5.2)$$

Esempio 5.3.1

Determinare la densità discreta associata alla variabile casuale discreta $\xi(\omega) = k$, essendo Ω lo spazio campione associato all'esperimento casuale, ripetuto n volte, avente i soli due risultati possibili 0 e 1, ed essendo k il numero di risultati pari a 1 compresi in ω .

Si tratta dello stesso esperimento casuale che ha condotto alla legge di Bernoulli (si veda l'Esempio 5.2.1), per il quale la probabilità di avere k risultati positivi su n ripetizioni dell'esperimento vale:

$$P_{k,n} = \binom{n}{k} p^k (1-p)^{n-k},$$

dove p è la probabilità che, per ogni esperimento casuale, il risultato sia 1. La densità discreta associata alla variabile casuale ξ vale pertanto:

$$p_\xi(x) = P(\{\xi = x\}) = B(n,p) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & \text{se } x = 0, 1, 2, \dots, n \\ 0 & \text{altrimenti} \end{cases}.$$

La distribuzione individuata da questa densità viene detta *legge binomiale* di parametri n e p . In particolare, per $n = 1$ ci si riduce al caso di una variabile casuale che può assumere solo i valori 0 (con probabilità $1 - p$) e 1 (con probabilità p): in questo caso si parla di *legge di Bernoulli* di parametro p . ♦

La legge binomiale, in definitiva, può essere utilizzata tutte le volte che si voglia determinare la probabilità che, in un esperimento con due possibili risultati, la ripetizione indipendente di n esperimenti dia al più k risultati positivi: tale probabilità vale evidentemente:

$$P(\{\xi \leq k\}) = \sum_{i=0}^k P(\{\xi = i\}) = \sum_{i=0}^k p_\xi(i) = \sum_{i=0}^k \binom{n}{i} p^i (1-p)^{n-i}.$$

Esempio 5.3.2

I bulloni prodotti da una ditta sono difettosi con una probabilità del 20%, e vengono messi in commercio in confezioni da 5 pezzi ciascuna. Quale è la probabilità che in una confezione vi sia al più un bullone difettoso?

Assumendo che il fatto che un bullone sia difettoso sia indipendente dagli altri bulloni, la probabilità richiesta vale $P(\{\xi \leq 1\})$, essendo ξ una variabile casuale discreta che conta il numero di bulloni difettosi. Sapendo che la probabilità che ha ogni bullone di essere difettoso vale $p = 0.2$, e che in ogni confezione vi sono 5 bulloni, la variabile casuale ξ avrà densità discreta data dalla legge di Bernoulli $B(5, 0.2)$, e la probabilità richiesta vale:

$$\begin{aligned} P(\{\xi \leq 1\}) &= P(\{\xi = 0\}) + P(\{\xi = 1\}) = \binom{5}{0} p^0 (1-p)^5 \\ &+ \binom{5}{1} p (1-p)^4 = 0.8^5 + 5 \cdot 0.2 \cdot 0.8^4 = 0.737. \end{aligned}$$

**Esempio 5.3.3**

Un'urna contiene 8 palline rosse e due bianche. Supponendo di estrarne 3, qual è la probabilità di averne estratte al più una bianca?

Si noti che questo tipo di esperimento casuale non può essere studiato con la legge binomiale, infatti l'estrazione delle 3 palline avviene in modo contemporaneo, e non rimettendo ogni volta nell'urna la pallina estratta. Abbiamo già visto negli Esempi 5.2.2 e 5.2.3 che in questo caso la probabilità segue la legge ipergeometrica, pertanto indicando con ξ la variabile casuale discreta che corrisponde al numero totale di palline bianche estratte, sarà:

$$P(\{\xi \leq 1\}) = P(\{\xi = 0\}) + P(\{\xi = 1\}) = \frac{\binom{2}{0} \binom{8}{3}}{\binom{10}{3}} + \frac{\binom{2}{1} \binom{8}{2}}{\binom{10}{3}} = 0.933.$$



Abbiamo visto negli Esempi precedenti come lo schema di Bernoulli sia utile quando si vuole stimare la probabilità di avere k successi qualora si ripeta lo stesso esperimento casuale un numero n finito di volte. È naturale, a questo punto, chiedersi come si modifichi questo ragionamento quando il numero di ripetizioni

dell'esperimento casuale diviene molto grande, al limite per $n \rightarrow +\infty$. Si può dimostrare che in tale limite la variabile casuale ξ corrispondente al numero di risultati positivi assume una densità discreta $p_\xi(x)$ cui corrisponde la *distribuzione di Poisson* di parametro λ :

$$p_\xi(x) = \begin{cases} e^{-\lambda} \frac{\lambda^x}{x!} & \text{se } x = 0, 1, 2, \dots \\ 0 & \text{altrimenti} \end{cases}.$$

Questa è un caso limite della distribuzione di Bernoulli nel senso che corrisponde al limite per $n \rightarrow +\infty$ di $B(n, \lambda/n)$. Ciò significa che se si è in presenza di una distribuzione di Bernoulli in cui n è molto grande, e il parametro p è molto piccolo, questa può essere approssimata da una distribuzione di Poisson.

In accordo con le considerazioni precedenti, alla distribuzione di Poisson di parametro λ corrisponde la probabilità:

$$P(\{\xi \leq n\}) = \sum_{k=0}^n e^{-\lambda} \frac{\lambda^k}{k!}.$$

Definizione 5.3.4 Sia ξ una variabile casuale (discreta o continua). Si chiama *funzione di ripartizione* o *funzione distribuzione cumulativa* di ξ la funzione $F_\xi: \mathbb{R} \rightarrow [0,1]$ definita da:

$$F_\xi(x) = P(\{\xi \leq x\}).$$

La funzione di ripartizione consente immediatamente di valutare la probabilità di un evento del tipo $\{x_1 < \xi \leq x_2\}$ con $x_1 \leq x_2$, infatti si ha:

$$\{\xi \leq x_2\} = \{\xi \leq x_1\} \cup \{x_1 < \xi \leq x_2\}$$

dove i due eventi $\{\xi \leq x_1\}$ e $\{x_1 < \xi \leq x_2\}$ sono disgiunti, pertanto:

$$P(\{\xi \leq x_2\}) = P(\{\xi \leq x_1\}) + P(\{x_1 < \xi \leq x_2\}),$$

da cui si ricava:

$$P(\{x_1 < \xi \leq x_2\}) = P(\{\xi \leq x_2\}) - P(\{\xi \leq x_1\}) = F_\xi(x_2) - F_\xi(x_1). \quad (5.3)$$

Teorema 5.3.2 Sia ξ una variabile casuale di funzione di ripartizione $F_\xi(x)$. Allora:

1. $F_\xi(x)$ è una funzione monotona non decrescente del suo argomento;

2. valgono le relazioni:

$$\lim_{x \rightarrow -\infty} F_\xi(x) = 0, \quad \lim_{x \rightarrow +\infty} F_\xi(x) = 1.$$

Nel caso particolare delle variabili aleatorie discrete, la funzione di distribuzione cumulativa assume una forma peculiare. Infatti, sia ξ una variabile casuale discreta che assume valori in un insieme numerabile $\{x_1, \dots, x_n, \dots\} \subset \mathbb{R}$, che possiamo assumere ordinato in senso crescente. Allora, in un qualunque intervallo $]x_i, x_{i+1}[$ si ha:

$$\{\xi \leq x\} = \{\xi \leq x_{i+1}\} \quad \forall x \in]x_i, x_{i+1}[,$$

e quindi la funzione di ripartizione risulta essere costante:

$$F_\xi(x) = F_\xi(x_{i+1}) \quad \forall x \in]x_i, x_{i+1}[.$$

In altri termini, la funzione di distribuzione cumulativa di una variabile casuale discreta è una funzione *costante a tratti*. Inoltre, nel caso delle variabili casuali discrete la conoscenza della funzione di ripartizione e della densità discreta sono equivalenti. Infatti dalla (5.2):

$$F_\xi(x) = \sum_{t \leq x} p_\xi(t),$$

viceversa, dati due valori consecutivi x_i e x_{i+1} (ordinati in senso crescente), per la (5.3) si ha:

$$F_\xi(x_{i+1}) - F_\xi(x_i) = P(\{x_i < \xi \leq x_{i+1}\}) = P(\{\xi = x_{i+1}\}) = p_\xi(x_{i+1}).$$

5.3.2 Speranza matematica

Consideriamo la una variabile casuale discreta ξ , caratterizzata da una densità discreta p_ξ e dai valori $\{x_1, x_2, \dots, x_n, \dots\}$.

Definizione 5.3.5 La variabile casuale discreta ξ ha *speranza matematica finita* se:

$$\sum_i |x_i| p_\xi(x_i) < +\infty.$$

In questo caso, si chiama *speranza matematica* (detta anche *media*, *valore medio* o *valore atteso*) di ξ la quantità:

$$E(\xi) = \sum_i x_i p_\xi(x_i).$$

Teorema 5.3.3 Sia ξ una variabile aleatoria discreta, e sia φ una funzione di variabile casuale. Allora $\eta = \varphi(\xi)$ ha speranza matematica finita se e solo se:

$$\sum_i |\varphi(x_i)| p_\xi(x_i) < +\infty,$$

e, in questo caso, si ha:

$$E(\eta) = \sum_i \varphi(x_i) p_\xi(x_i).$$

Teorema 5.3.4 (Linearità della media) Siano ξ, η due variabili casuali discrete aventi speranza matematica finita, e siano $a, b \in \mathbb{R}$. Allora:

$$E(a\xi + b\eta) = aE(\xi) + bE(\eta).$$

Teorema 5.3.5 Siano ξ, η due variabili casuali discrete indipendenti aventi speranza matematica finita. Allora anche $\xi\eta$ è una variabile casuale con speranza matematica finita, e:

$$E(\xi\eta) = E(\xi)E(\eta).$$

Esempio 5.3.4

Si consideri una variabile casuale ξ binomiale, ovvero caratterizzata dalla legge binomiale $B(n, p)$ di parametro p descritta nell'Esempio 5.3.1. Grazie alla Definizione 5.3.5:

$$E(\xi) = \sum_{k=0}^n k p_\xi(k) = \sum_{k=0}^n k \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=1}^n k \binom{n}{k} p^k (1-p)^{n-k}.$$

Questa espressione può essere esplicitata ricordando lo sviluppo del binomio:

$$(a+b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k} \quad a, b \in \mathbb{R}.$$

Infatti per $k > 0$:

$$k \binom{n}{k} = k \frac{n!}{k!(n-k)!} = n \frac{(n-1)!}{(k-1)!(n-k)!} = n \binom{n-1}{k-1},$$

per cui, ponendo $m = k - 1$:

$$\begin{aligned} E(\xi) &= np \sum_{k=1}^n \binom{n-1}{k-1} p^{k-1} (1-p)^{n-k} = np \sum_{m=0}^{n-1} \binom{n-1}{m} p^m (1-p)^{n-1-m} \\ &= np [p + (1-p)]^{n-1} = np. \end{aligned}$$

In particolare, per la legge di Bernoulli $B(1,p)$ si ha:

$$E(\xi) = \sum_{k=0}^1 kp_{\xi}(k) = 0(1-p) + 1p = p.$$

◆

Esempio 5.3.5

Sia ξ una variabile casuale discreta con distribuzione di Poisson di parametro λ . Il relativo valore medio è (si pone $q = k - 1$):

$$\begin{aligned} E(\xi) &= \sum_{k=0}^{+\infty} kp_{\xi}(k) = \sum_{k=0}^{+\infty} ke^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=1}^{+\infty} \frac{\lambda^k}{(k-1)!} \\ &= e^{-\lambda} \lambda \sum_{k=1}^{+\infty} \frac{\lambda^{k-1}}{(k-1)!} = e^{-\lambda} \lambda \sum_{q=0}^{+\infty} \frac{\lambda^q}{q!} = e^{-\lambda} \lambda e^{\lambda} = \lambda. \end{aligned}$$

◆

Una volta definito il concetto di valore medio di una variabile aleatoria, questo viene a sua volta utilizzato per introdurre altre grandezze importanti nella caratterizzazione delle variabili casuali.

Definizione 5.3.6 Sia ξ una variabile casuale discreta. Fissato $k = 1, 2, \dots$, si dice che ξ ha *momento di ordine k* finito se la variabile casuale ξ^k ha speranza matematica finita. In questo caso, si chiama momento di ordine k di ξ la quantità $E(\xi^k)$. Analogamente, se ξ ha speranza matematica finita $\mu = E(\xi)$, se la variabile aleatoria $[\xi - E(\xi)]^k = (\xi - \mu)^k$ ha speranza matematica finita, diremo che ξ ha *momento centrale di ordine k* finito, e chiameremo momento centrale di ordine k la quantità $E([\xi - E(\xi)]^k)$.

Grazie al Teorema 5.3.3, si ricava immediatamente:

$$\begin{aligned} E(\xi^k) &= \sum_i x_i^k p_{\xi}(x_i), \\ E([\xi - E(\xi)]^k) &= \sum_i (x_i - \mu)^k p_{\xi}(x_i), \quad \mu = E(\xi). \end{aligned}$$

Teorema 5.3.6 Siano ξ e η due variabili aleatorie discrete con momento di ordine k finito. Allora:

1. le due variabili casuali hanno momento di ordine r finito per ogni $r \leq k$;
2. anche la variabile aleatoria $\xi + \eta$ ha momento di ordine k finito.

Di particolare importanza pratica è il momento centrale di ordine 2, di solito chiamato *varianza* $\text{var}(\xi)$:

$$\text{var}(\xi) = \sigma^2 = E([\xi - E(\xi)]^2).$$

Per la linearità della media, posto $\mu = E(\xi)$ si ha:

$$\text{var}(\xi) = \sigma^2 = E(\xi^2) - \mu^2 = E(\xi^2) - E(\xi)^2,$$

inoltre si verifica facilmente che, fissato un qualunque $a \in \mathbb{R}$:

$$\text{var}(a\xi) = a^2 \text{var}(\xi), \quad \text{var}(\xi + a) = \text{var}(\xi).$$

La varianza di una variabile casuale è una misura della dispersione di ξ rispetto al suo valore medio. Infatti, se ξ può assumere solo il valore $a \in \mathbb{R}$ con probabilità 1, si ha $\mu = a$ e $\text{var}(\xi) = 0$. In particolare, la grandezza $\sigma = \sqrt{\text{var}(\xi)}$ viene detta *deviazione standard* della variabile aleatoria ξ .

Esempio 5.3.6

Si consideri una variabile casuale ξ binomiale. Dalla definizione si ricava immediatamente:

$$E(\xi^2) = \sum_{k=0}^n k^2 p_\xi(k) = \sum_{k=0}^n k^2 \binom{n}{k} p^k (1-p)^{n-k}$$

e quindi, con considerazioni analoghe a quelle condotte nell'Esempio 5.3.4 (si pone nuovamente $m = k - 1$):

$$E(\xi^2) = np \sum_{k=1}^n k \binom{n-1}{k-1} p^{k-1} (1-p)^{n-k} = np \sum_{m=0}^{n-1} (m+1) \binom{n-1}{m} p^m (1-p)^{n-1-m}.$$

La sommatoria si può decomporre nella somma di due addendi:

$$E(\xi^2) = np \left[\sum_{m=0}^{n-1} m \binom{n-1}{m} p^m (1-p)^{n-1-m} + \sum_{m=0}^{n-1} \binom{n-1}{m} p^m (1-p)^{n-1-m} \right],$$

dove:

$$\sum_{m=0}^{n-1} \binom{n-1}{m} p^m (1-p)^{n-1-m} = [p + (1-p)]^{n-1} = 1,$$

mentre, per il primo addendo, si osserva che (per $m > 0$):

$$m \binom{n-1}{m} = (n-1) \binom{n-2}{m-1},$$

e quindi, posto $q = m - 1$:

$$\begin{aligned} \sum_{m=0}^{n-1} m \binom{n-1}{m} p^m (1-p)^{n-1-m} &= (n-1)p \sum_{m=1}^{n-1} \binom{n-2}{m-1} p^{m-1} (1-p)^{n-1-m} \\ &= (n-1)p \sum_{q=0}^{n-2} \binom{n-2}{q} p^q (1-p)^{n-2-q} \\ &= (n-1)p [p + (1-p)]^{n-2} = (n-1)p. \end{aligned}$$

In definitiva:

$$E(\xi^2) = np[(n-1)p + 1] = n(n-1)p^2 + np,$$

mentre la varianza vale:

$$\text{var}(\xi) = E(\xi^2) - E(\xi)^2 = n(n-1)p^2 + np - n^2p^2 = np(1-p).$$

In particolare, per la legge di Bernoulli $B(1,p)$ si ha:

$$E(\xi^2) = p, \quad \text{var}(\xi) = p(1-p).$$

◆

Esempio 5.3.7

Sia ξ una variabile casuale discreta con distribuzione di Poisson di parametro λ . Nell'Esempio 5.3.5 abbiamo visto che $\mu = E(\xi) = \lambda$, inoltre (posto $q = k - 1$):

$$\begin{aligned} E(\xi^2) &= e^{-\lambda} \sum_{k=0}^{+\infty} k^2 \frac{\lambda^k}{k!} = e^{-\lambda} \lambda \sum_{k=1}^{+\infty} k \frac{\lambda^{k-1}}{(k-1)!} \\ &= e^{-\lambda} \lambda \sum_{q=0}^{+\infty} (q+1) \frac{\lambda^q}{q!} = \lambda E(\xi + 1) = \lambda(\lambda + 1). \end{aligned}$$

Pertanto:

$$\text{var}(\xi) = E(\xi^2) - E(\xi)^2 = \lambda(\lambda + 1) - \lambda^2 = \lambda,$$

e quindi per una variabile aleatoria con distribuzione di Poisson, il parametro coincide sia con il valore medio sia con la varianza. ◆

Si considerino, ora, due variabili casuali discrete ξ e η aventi momento del secondo ordine finito. Allora, per il Teorema 5.3.6, anche $\xi + \eta$ ha momento del secondo ordine finito, ed è:

$$\begin{aligned}\text{var}(\xi + \eta) &= E[(\xi + \eta)^2 - E(\xi + \eta)^2] = E\{[(\xi - E(\xi)) + (\eta - E(\eta))]^2\} \\ &= E[(\xi - E(\xi))^2] + E[(\eta - E(\eta))^2] + 2E[(\xi - E(\xi))(\eta - E(\eta))],\end{aligned}$$

infatti:

$$E[(\xi + \eta)^2 - E(\xi + \eta)^2] = E(\xi^2) + E(\eta^2) + 2E(\xi\eta) - E(\xi)^2 - E(\eta)^2 - 2E(\xi)E(\eta),$$

mentre da:

$$\begin{aligned}[(\xi - E(\xi)) + (\eta - E(\eta))]^2 &= \xi^2 + E(\xi)^2 - 2\xi E(\xi) + \eta^2 + E(\eta)^2 - 2\eta E(\eta) \\ &\quad + 2\xi\eta + 2E(\xi)E(\eta) - 2\xi E(\eta) - 2\eta E(\xi)\end{aligned}$$

si ricava facilmente:

$$E\{[(\xi - E(\xi)) + (\eta - E(\eta))]^2\} = E(\xi^2) + E(\eta^2) + 2E(\xi\eta) - E(\xi)^2 - E(\eta)^2 - 2E(\xi)E(\eta).$$

Si definisce *covarianza* di ξ e η la quantità:

$$\text{cov}(\xi, \eta) = E[(\xi - E(\xi))(\eta - E(\eta))] = E(\xi\eta) - E(\xi)E(\eta).$$

Grazie al Teorema 5.3.5, se ξ e η sono indipendenti, si ha $\text{cov}(\xi, \eta) = 0$.

Queste considerazioni possono essere generalizzate al caso di n variabili aleatorie:

Teorema 5.3.7 Siano $\xi_1, \xi_2, \dots, \xi_n$ n variabili casuali discrete con momento del secondo ordine finito. Allora:

$$\text{var}(\xi_1 + \dots + \xi_n) = \sum_{k=1}^n \text{var}(\xi_k) + \sum_{\substack{k, l=1 \\ k \neq l}}^n \text{cov}(\xi_k, \xi_l).$$

Se le n variabili casuali sono anche indipendenti, allora:

$$\text{var}(\xi_1 + \dots + \xi_n) = \sum_{k=1}^n \text{var}(\xi_k).$$

Fino ad ora ci siamo occupati della definizione e del calcolo della media e della varianza di una variabile casuale discreta ξ , che abbiamo visto dipendere esclusivamente dalla densità discreta di ξ . In pratica, spesso, accade di avere a disposizione una serie di valori numerici corrispondenti al risultato di numerose ripetizioni di un esperimento casuale, ma senza avere, a priori, alcuna informazione sulla forma della densità discreta corrispondente. In altri termini, si considerino n campioni x_1, x_2, \dots, x_n , che si assumono essere corrispondenti ad altrettanti valori assunti da una variabile casuale ξ . Il problema che vogliamo porci è quello di costruire delle approssimazioni, il più possibile buone, di $E(\xi)$ e di $\text{var}(\xi)$: si parla, in questo caso, di problemi di *statistica matematica*. Una analisi approfondita consente di dimostrare come una buona stima per i due parametri in esame, tanto migliore quanto più è elevato il numero n di campioni, è data da:

$$E(\xi) \approx \bar{x} = \frac{1}{n} \sum_{k=1}^n x_k, \quad \text{var}(\xi) \approx s^2 = \frac{1}{n-1} \sum_{k=1}^n (x_k - \bar{x})^2.$$

Esempio 5.3.8

Viene osservato il numero di figli maschi in 200 famiglie di 4 figli, con il risultato:

n	0	1	2	3	4
N_n	26	49	54	45	26

essendo N_n il numero di famiglie che hanno n figli maschi. Per stimare il valore medio della distribuzione dei figli maschi nelle 200 famiglie, si usa:

$$\bar{n} = \frac{1}{200} \sum_{n=0}^4 nN_n = \frac{0 \cdot 26 + 1 \cdot 49 + 2 \cdot 54 + 3 \cdot 45 + 4 \cdot 26}{200} = 1.98,$$

mentre la varianza viene stimata secondo:

$$s^2 = \frac{1}{199} \sum_{n=0}^4 N_n (n - \bar{n})^2 = \frac{26(0 - 1.98)^2 + 49(1 - 1.98)^2 + 54(2 - 1.98)^2}{199} + \frac{45(3 - 1.98)^2 + 26(4 - 1.98)^2}{199} = 1.517.$$



5.4 Variabili casuali continue

Le variabili casuali continue, già introdotte nella Sezione 5.3, sono caratterizzate dall'assumere valori in un sottoinsieme di \mathbb{R} avente la cardinalità del continuo. Anche in questo caso, assegnata una variabile aleatoria (continua) ξ è possibile definire la *funzione di ripartizione*, o *funzione di distribuzione cumulativa* secondo:

$$F_\xi(x) = P(\{\xi \leq x\}),$$

che gode di proprietà analoghe alla corrispondente funzione definita per le variabili casuali discrete. In particolare, se $a, b \in \mathbb{R}$ con $a \leq b$:

$$P(\{a < \xi \leq b\}) = P(\{\xi \leq b\}) - P(\{\xi \leq a\}) = F_\xi(b) - F_\xi(a),$$

e vale il Teorema 5.3.2 sulla monotonia e limitatezza di F_ξ .

Teorema 5.4.1 Una variabile casuale ξ è continua se e solo se la sua funzione di distribuzione cumulativa $F_\xi(x)$ è continua, e in questo caso si ha:

$$P(\{\xi = x\}) = 0.$$

Naturalmente, è immediato verificare che per una variabile aleatoria continua si ha:

$$P(\{a < \xi < b\}) = P(\{a < \xi \leq b\}) = P(\{a \leq \xi < b\}) = P(\{a \leq \xi \leq b\}).$$

Anche il concetto di densità discreta può essere opportunamente generalizzato al caso delle variabili aleatorie continue:

Definizione 5.4.1 Una funzione $f: \mathbb{R} \rightarrow \mathbb{R}$ viene detta una *densità* se e solo se è una funzione non negativa, integrabile su \mathbb{R} e tale che:

$$\int_{-\infty}^{+\infty} f(x) \, dx = 1.$$

Definizione 5.4.2 Sia ξ una variabile casuale continua di funzione di distribuzione cumulativa $F_\xi(x)$. Si dice che ξ ha *densità di probabilità* f_ξ se:

$$F_\xi(x) = \int_{-\infty}^x f_\xi(t) \, dt.$$

Dalla definizione è immediato verificare che:

$$P(\{a < \xi \leq b\}) = F_\xi(b) - F_\xi(a) = \int_a^b f_\xi(x) dx.$$

Le variabili casuali continue che hanno una densità di probabilità vengono dette *assolutamente continue*.

Dalla definizione, si deduce immediatamente che, per una variabile aleatoria assolutamente continua, le regioni in cui f_ξ assume valore elevato sono quelle in cui ξ prende valori con probabilità elevata. Inoltre, la conoscenza della densità di probabilità consente di risalire, per integrazione, alla funzione di ripartizione. Supponiamo, ora, di conoscere F_ξ e di voler determinare la densità di probabilità. Poiché F_ξ è la funzione integrale di f_ξ , se F_ξ è derivabile con derivata continua in \mathbb{R} (tranne al più in un numero finito di punti), allora F_ξ è la funzione integrale della sua derivata, e quindi si ha:

$$f_\xi(x) = \frac{dF_\xi}{dx},$$

pertanto la derivata prima della funzione di distribuzione cumulativa può essere usata come densità di probabilità per ξ .

Esempio 5.4.1

Si consideri una variabile casuale assolutamente continua caratterizzata dalla funzione di distribuzione cumulativa:

$$F_\xi(x) = \begin{cases} 0 & \text{se } x < 0 \\ x & \text{se } 0 \leq x \leq 1 \\ 1 & \text{se } x > 1 \end{cases}.$$

Questa è derivabile con derivata continua, tranne che nei punti $x = 0$ e $x = 1$, pertanto la densità di probabilità associata a F_ξ è:

$$f_\xi(x) = \frac{dF_\xi}{dx} = \begin{cases} 1 & \text{se } 0 \leq x \leq 1 \\ 0 & \text{altrimenti} \end{cases},$$

cui corrisponde una variabile aleatoria ξ per la quale:

$$P(\{a \leq \xi \leq b\}) = \int_a^b f_\xi(x) dx = b - a$$

qualunque siano $a, b \in [0, 1]$. In altri termini, la probabilità che ξ assuma valori in un sottointervallo di $[0, 1]$ dipende solo dall'ampiezza del sottointervallo stesso: la variabile casuale così definita viene detta *uniforme* in $[0, 1]$. ♦

Esempio 5.4.2

Si consideri una variabile casuale assolutamente continua caratterizzata dalla funzione di distribuzione cumulativa:

$$F_{\xi}(x) = \begin{cases} 1 - e^{-\lambda x} & \text{se } x > 0 \\ 0 & \text{altrimenti} \end{cases}.$$

Anche in questo caso, si tratta di una funzione derivabile ovunque tranne che per $x = 0$, quindi la corrispondente densità di probabilità può essere valutata per derivazione:

$$f_{\xi}(x) = \begin{cases} \lambda e^{-\lambda x} & \text{se } x > 0 \\ 0 & \text{altrimenti} \end{cases}.$$

Questa densità di probabilità viene detta *densità esponenziale di parametro λ* . ♦

Si considerino, ora, due variabili casuali continue ξ e η . Per definizione, sia $\{\xi \leq x\}$ sia $\{\eta \leq y\}$ sono degli eventi, quindi è un evento anche $\{\xi \leq x, \eta \leq y\} = \{\xi \leq x\} \cap \{\eta \leq y\}$, che si può esprimere come l'evento:

$$A_{x,y} = \{(u,v) \in \mathbb{R}^2 : u \leq x, v \leq y\}.$$

Si indica con $F_{\xi,\eta}(x,y)$ la *funzione di ripartizione congiunta* di ξ e η , definita da:

$$F_{\xi,\eta}(x,y) = P(A_{x,y}) = P(\{\xi \leq x, \eta \leq y\}).$$

Si dirà inoltre che ξ e η hanno *densità di probabilità congiunta* $f_{\xi,\eta}(x,y)$ se esiste una funzione $f_{\xi,\eta}$ integrabile e non negativa tale che:

$$F_{\xi,\eta}(x,y) = \int_{A_{x,y}} f_{\xi,\eta}(u,v) \, dudv = \int_{-\infty}^x \int_{-\infty}^y f_{\xi,\eta}(u,v) \, dudv.$$

Le funzioni di ripartizione (risp. densità di probabilità) $F_{\xi}(x)$ e $F_{\eta}(y)$ (risp. $f_{\xi}(x)$ e $f_{\eta}(y)$) associate alle due variabili casuali che costituiscono la coppia ordinata (ξ, η) vengono chiamate funzioni di ripartizione (risp. densità di probabilità) *marginali* di ξ e η .

Teorema 5.4.2 Siano ξ e η due variabili casuali continue caratterizzate dalla densità di probabilità congiunta $f_{\xi,\eta}(x,y)$. Allora:

$$\int_{\mathbb{R}^2} f_{\xi,\eta}(x,y) \, dx dy = 1.$$

Teorema 5.4.3 Siano ξ e η due variabili casuali continue caratterizzate dalle funzioni di ripartizione e dalle densità di probabilità marginali $F_\xi(x)$, $F_\eta(y)$, $f_\xi(x)$ e $f_\eta(y)$, rispettivamente. Siano inoltre $F_{\xi,\eta}(x,y)$ e $f_{\xi,\eta}(x,y)$, rispettivamente, la funzione di distribuzione cumulativa e la densità di probabilità congiunta. Allora:

$$F_\xi(x) = \lim_{y \rightarrow +\infty} F_{\xi,\eta}(x,y), \quad F_\eta(y) = \lim_{x \rightarrow +\infty} F_{\xi,\eta}(x,y),$$

$$f_\xi(x) = \int_{-\infty}^{+\infty} f_{\xi,\eta}(x,y) dy, \quad f_\eta(y) = \int_{-\infty}^{+\infty} f_{\xi,\eta}(x,y) dx.$$

Definizione 5.4.3 Siano ξ_1, \dots, ξ_n n variabili casuali continue. Esse vengono dette *indipendenti* se per qualunque scelta di $a_1, b_1, \dots, a_n, b_n \in \mathbb{R}$ (tali che $a_i \leq b_i$ per $i = 1, \dots, n$) si ha:

$$P(\{a_1 \leq \xi_1 \leq b_1, \dots, a_n \leq \xi_n \leq b_n\}) = P(\{a_1 \leq \xi_1 \leq b_1\}) \dots P(\{a_n \leq \xi_n \leq b_n\}).$$

Se le variabili aleatorie continue sono in un numero infinito, la definizione di indipendenza si estende richiedendo l'indipendenza di qualunque loro sottoinsieme finito.

Teorema 5.4.4 Siano ξ e η due variabili casuali continue di densità di probabilità congiunta $f_{\xi,\eta}(x,y)$ e densità marginali $f_\xi(x)$ e $f_\eta(y)$. Esse sono indipendenti se e solo se $f_{\xi,\eta}(x,y) = f_\xi(x)f_\eta(y)$ in tutto \mathbb{R}^2 , tranne al più in un insieme di punti $A \subset \mathbb{R}^2$ tale che $m(A) = 0$.

Come già discusso nella Sezione 5.3, date una variabile casuale continua ξ e una funzione $\varphi: \mathbb{R} \rightarrow \mathbb{R}$ a valori reali, se questa è sufficientemente regolare anche $\varphi(\xi)$ è una variabile casuale.

Teorema 5.4.5 Siano ξ_1, \dots, ξ_n n variabili casuali continue indipendenti, e siano $\varphi_1, \dots, \varphi_n$ delle applicazioni sufficientemente regolari perché $\eta_1 = \varphi_1(\xi_1), \dots, \eta_n = \varphi_n(\xi_n)$ siano variabili casuali. Allora, queste sono indipendenti.

5.4.1 Funzioni di variabili casuali assolutamente continue

Consideriamo una variabile casuale assolutamente continua ξ di distribuzione cumulativa $F_\xi(x)$ e densità $f_\xi(x) = F'_\xi(x)$, e una funzione $\varphi: \mathbb{R} \rightarrow \mathbb{R}$ sufficientemente regolare perché anche $\eta = \varphi(\xi)$ sia una variabile casuale. La funzione di distribuzione cumulativa di η , per definizione, è data da:

$$F_\eta(y) = P(\{\eta \leq y\}) = P(\{\varphi(\xi) \leq y\}).$$

Posto:

$$D_y = \{x \in \mathbb{R} : \varphi(x) \leq y\},$$

l'evento $\{\varphi(\xi) \leq y\}$ coincide con l'evento $\{\xi \in D_y\}$ e quindi:

$$F_\eta(y) = P(\{\eta \leq y\}) = P(\{\xi \in D_y\}) = \int_{D_y} f_\xi(x) dx. \quad (5.4)$$

Infine, se la funzione di ripartizione $F_\eta(y)$ così calcolata risulta essere derivabile con derivata continua tranne al più un numero finito di punti, si valuta:

$$f_\eta(y) = \frac{dF_\eta}{dy}. \quad (5.5)$$

Esempio 5.4.3

Sia $\eta = \xi^2$ una variabile casuale ottenuta trasformando la variabile aleatoria assolutamente continua ξ attraverso la funzione $\varphi(x) = x^2$. Allora, fissato y la condizione $\varphi(x) = x^2 \leq y$ è soddisfatta per $-\sqrt{y} \leq x \leq \sqrt{y}$, per cui:

$$D_y = \{x \in \mathbb{R} : -\sqrt{y} \leq x \leq \sqrt{y}\}.$$

Applicando la (5.4) si ha quindi:

$$F_\eta(y) = \int_{D_y} f_\xi(x) dx = \int_{-\sqrt{y}}^{\sqrt{y}} f_\xi(x) dx = F_\xi(\sqrt{y}) - F_\xi(-\sqrt{y}).$$

Infine, usando la (5.5):

$$f_\eta(y) = \frac{dF_\eta}{dy} = \frac{f_\xi(\sqrt{y}) + f_\xi(-\sqrt{y})}{2\sqrt{y}}.$$

In particolare, se ad esempio ξ ha una densità uniforme, trasformandola con φ si ottiene una variabile casuale con densità del tipo $1/\sqrt{y}$. \blacklozenge

Esempio 5.4.4

Sia $\eta = a\xi + b$ una variabile casuale ottenuta trasformando la variabile aleatoria assolutamente continua ξ attraverso la funzione lineare $\varphi(x) = ax + b$, $a, b \in \mathbb{R}$. Allora, fissato y , per $a > 0$ la condizione $\varphi(x) = ax + b \leq y$ è soddisfatta per $-\infty \leq x \leq (y - b)/a$ (si veda la Fig. 5.1 a sinistra), per cui:

$$D_y = \{x \in \mathbb{R} : -\infty \leq x \leq (y - b)/a\}.$$

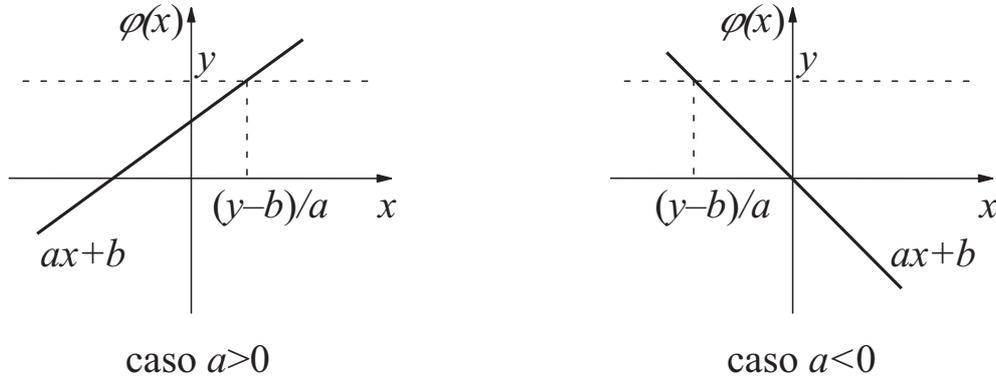


Figura 5.1. Rappresentazione grafica dell'insieme D_y per la trasformazione lineare dell'Esempio 5.4.4 con $a > 0$ (a sinistra) e $a < 0$ (a destra).

Applicando la (5.4) si ha quindi:

$$F_\eta(y) = \int_{D_y} f_\xi(x) \, dx = \int_{-\infty}^{(y-b)/a} f_\xi(x) \, dx = F_\xi\left(\frac{y-b}{a}\right) - F_\xi(-\infty) = F_\xi\left(\frac{y-b}{a}\right).$$

Infine, usando la (5.5):

$$f_\eta(y) = \frac{dF_\eta}{dy} = \frac{1}{a} f_\xi\left(\frac{y-b}{a}\right).$$

Analogamente, dalla parte destra di Fig. 5.1, si deduce che per $a < 0$ la condizione $\varphi(x) = ax + b \leq y$ è soddisfatta per $(y-b)/a \leq x \leq +\infty$ e quindi:

$$D_y = \{x \in \mathbb{R} : (y-b)/a \leq x \leq +\infty\}.$$

Pertanto, si ha facilmente:

$$F_\eta(y) = \int_{D_y} f_\xi(x) \, dx = \int_{(y-b)/a}^{+\infty} f_\xi(x) \, dx = F_\xi(+\infty) - F_\xi\left(\frac{y-b}{a}\right) = 1 - F_\xi\left(\frac{y-b}{a}\right),$$

$$f_\eta(y) = \frac{dF_\eta}{dy} = -\frac{1}{a} f_\xi\left(\frac{y-b}{a}\right).$$

In definitiva, riunendo i due risultati precedenti si ha, indipendentemente dal segno di a :

$$f_\eta(y) = \frac{1}{|a|} f_\xi\left(\frac{y-b}{a}\right).$$

◆

5.4.2 Variabili casuali gaussiane

Si consideri la funzione di variabile reale:

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2},$$

detta *gaussiana*. Si può dimostrare che, sebbene l'integrale di $\exp(-x^2/2)$ non sia esprimibile in forma chiusa, vale il risultato:

$$\int_{-\infty}^{+\infty} e^{-x^2/2} dx = \sqrt{2\pi}. \quad (5.6)$$

Pertanto, in accordo alla Definizione 5.4.1, la funzione f è una densità. Una variabile casuale ξ caratterizzata dalla densità di probabilità:

$$f_\xi(x) = N(0,1) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

viene detta *variabile casuale gaussiana* o *normale* di parametri 0 e 1. Supponiamo, ora, di trasformare linearmente ξ nella variabile aleatoria $\eta = \sigma\xi + \mu$, con $\sigma > 0$. Grazie all'Esempio 5.4.4, si ha:

$$f_\eta(x) = \frac{1}{\sigma} f_\xi\left(\frac{x-\mu}{\sigma}\right) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] = N(\mu,\sigma).$$

Si parla, in questo caso, di variabile casuale gaussiana (o normale) di parametri μ e σ . La Fig. 5.2 mostra un confronto tra le due densità $N(0,1)$ e $N(0,\sqrt{2})$.

Assegnato un evento $\{\eta \leq x\}$, si ha:

$$P(\{\eta \leq x\}) = F_\eta(x) = \int_{-\infty}^x f_\eta(x) dx = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] dx.$$

Posto:

$$u = \frac{x-\mu}{\sqrt{2}\sigma} \implies du = \frac{1}{\sqrt{2}\sigma} dx,$$

l'integrale si trasforma in:

$$F_\eta(x) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{(x-\mu)/(\sqrt{2}\sigma)} e^{-u^2} du,$$

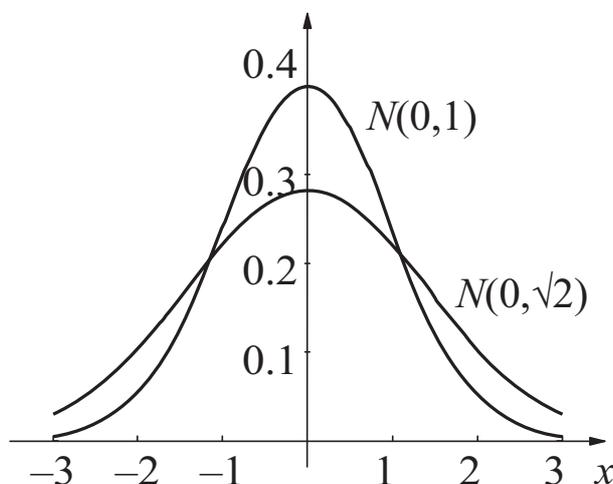


Figura 5.2. Confronto tra le densità di probabilità gaussiane $N(0,1)$ e $N(0,\sqrt{2})$.

dove l'intervallo di integrazione $] -\infty, (x - \mu)/(\sqrt{2}\sigma)]$ può essere decomposto in $] -\infty, 0] \cup] 0, (x - \mu)/(\sqrt{2}\sigma)]$. Grazie alla parità della funzione gaussiana, si ha immediatamente:

$$\int_{-\infty}^0 e^{-u^2} du = \frac{1}{2} \int_{-\infty}^{+\infty} e^{-u^2} du = \frac{\sqrt{\pi}}{2},$$

infatti, effettuando nella (5.6) il cambiamento di variabile $v = x/\sqrt{2} \implies dv = dx/\sqrt{2}$, si ha:

$$\int_{-\infty}^{+\infty} e^{-v^2} dv = \frac{\sqrt{2\pi}}{\sqrt{2}} = \sqrt{\pi}.$$

Si ha così

$$F_{\eta}(x) = \frac{1}{\sqrt{\pi}} \left[\frac{\sqrt{\pi}}{2} + \int_0^{(x-\mu)/(\sqrt{2}\sigma)} e^{-u^2} du, \right] = \frac{1}{2} \left[1 + \operatorname{erf} \left(\frac{x - \mu}{\sqrt{2}\sigma} \right) \right],$$

dove si è definita la *funzione di errore*:

$$\operatorname{erf}(y) = \frac{\sqrt{\pi}}{2} \int_0^y e^{-u^2} du.$$

Evidentemente, $F_{\eta}(x)$ non ammette espressione in forma chiusa. Si può, comunque, fare uso di tabelle che raccolgono un certo insieme di valori della funzione di distribuzione cumulativa per la variabile gaussiana di parametri 0 e 1, ovvero la variabile aleatoria ξ . Infatti, grazie alla trasformazione lineare $\eta = \sigma\xi + \mu$, la valutazione di

$F_\eta(x)$ può essere ricondotta a $F_\xi(x)$ grazie alla relazione (si confronti con l'Esempio 5.4.4):

$$F_\eta(x) = F_\xi\left(\frac{x - \mu}{\sigma}\right).$$

Alcuni valori di $F_\xi(x)$, per $x \geq 0$ sono riportati in Tabella 5.1. Nel caso sia necessario valutare la funzione di ripartizione in un argomento $-x < 0$, si può fare uso della parità della funzione integranda $f_\xi(x)$ per ricavare:

$$F_\xi(-x) = 1 - F_\xi(x) \quad \text{per } x \geq 0.$$

Esempio 5.4.5

Si consideri una variabile casuale η gaussiana di parametri $\mu = 3$ e $\sigma = 5$. Calcolare la probabilità che $-1 \leq \eta \leq 7$ e che $\eta > 8$.

Si ha immediatamente:

$$\begin{aligned} P(\{-1 \leq \eta \leq 7\}) &= F_\eta(7) - F_\eta(-1), \\ P(\{\eta > 8\}) &= 1 - P(\{\eta \leq 8\}) = 1 - F_\eta(8). \end{aligned}$$

Grazie all'Esempio 5.4.4, si ha:

$$\begin{aligned} F_\eta(7) &= F_\xi\left(\frac{7-3}{5}\right) = F_\xi(0.80), \\ F_\eta(-1) &= F_\xi\left(\frac{-1-3}{5}\right) = F_\xi(-0.80) = 1 - F_\xi(0.80), \\ F_\eta(8) &= F_\xi\left(\frac{8-3}{5}\right) = F_\xi(1.00). \end{aligned}$$

Dalla Tabella 5.1 si ha:

$$F_\xi(0.80) = 0.78814, \quad F_\xi(1.00) = 0.84134,$$

e quindi:

$$\begin{aligned} P(\{-1 \leq \eta \leq 7\}) &= F_\xi(0.80) - [1 - F_\xi(0.80)] = 0.57628, \\ P(\{\eta > 8\}) &= 1 - F_\xi(1.00) = 0.15866. \end{aligned}$$



x	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.50000	.50399	.50798	.51197	.51595	.51994	.52392	.52790	.53188	.53586
0.1	.53983	.54380	.54776	.55172	.55567	.55962	.56356	.56750	.57142	.57535
0.2	.57926	.58317	.58706	.59095	.59483	.59871	.60257	.60642	.61026	.61409
0.3	.61791	.62172	.62552	.62930	.63307	.63683	.64058	.64431	.64803	.65173
0.4	.65542	.65910	.66276	.66640	.67003	.67364	.67724	.68082	.68439	.68793
0.5	.69146	.69497	.69847	.70194	.70540	.70884	.71226	.71566	.71904	.72240
0.6	.72575	.72907	.73237	.73565	.73891	.74215	.74537	.74857	.75175	.75490
0.7	.75804	.76115	.76424	.76731	.77035	.77337	.77637	.77935	.78230	.78524
0.8	.78814	.79103	.79389	.79673	.79955	.80234	.80511	.80785	.81057	.81327
0.9	.81594	.81859	.82121	.82381	.82639	.82894	.83147	.83398	.83646	.83891
1.0	.84134	.84375	.84614	.84850	.85083	.85314	.85543	.85769	.85993	.86214
1.1	.86433	.86650	.86864	.87076	.87286	.87493	.87698	.87900	.88100	.88298
1.2	.88493	.88686	.88877	.89065	.89251	.89435	.89617	.89796	.89973	.90147
1.3	.90320	.90490	.90658	.90824	.90988	.91149	.91309	.91466	.91621	.91774
1.4	.91924	.92073	.92220	.92364	.92507	.92647	.92786	.92922	.93056	.93189
1.5	.93319	.93448	.93574	.93699	.93822	.93943	.94062	.94179	.94295	.94408
1.6	.94520	.94630	.94738	.94845	.94950	.95053	.95154	.95254	.95352	.95449
1.7	.95543	.95637	.95728	.95819	.95907	.95994	.96080	.9616	.96246	.96327
1.8	.96407	.96485	.96562	.96638	.96712	.96784	.96856	.96926	.96995	.97062
1.9	.97128	.97193	.97257	.97320	.97381	.97441	.97500	.97558	.97615	.97670
2.0	.97725	.97778	.97831	.97882	.97933	.97982	.98030	.98077	.98124	.98169
2.1	.98214	.98257	.98300	.98341	.98382	.98422	.98461	.98500	.98537	.98574
2.2	.98610	.98645	.98679	.98713	.98745	.98778	.98809	.98840	.98870	.98899
2.3	.98928	.98956	.98983	.99010	.99036	.99061	.99086	.99111	.99134	.99158
2.4	.99180	.99202	.99224	.99245	.99266	.99286	.99305	.99324	.99343	.99361
2.5	.99379	.99396	.99413	.99430	.99446	.99461	.99477	.99492	.99506	.99520
2.6	.99534	.99547	.99560	.99573	.99585	.99598	.99609	.99621	.99632	.99643
2.7	.99653	.99664	.99674	.99683	.99693	.99702	.99711	.99720	.99728	.99736
2.8	.99745	.99752	.99760	.99767	.99774	.99781	.99788	.99795	.99801	.99807
2.9	.99813	.99819	.99825	.99831	.99836	.99841	.99846	.99851	.99856	.99861

Tabella 5.1. Valori numerici della probabilità $P(\{\xi \leq x\}) = F_\xi(x)$ per una variabile casuale gaussiana con densità $N(0,1)$. La prima colonna indica il valore di x fino alla prima cifra decimale, mentre la colonna corrisponde alla seconda cifra decimale. Ad esempio, $P(\{\xi \leq 1.74\})$ si legge in corrispondenza della riga 1.7 e della colonna .04: $P(\{\xi \leq 1.74\}) = .95907$.

5.4.3 Speranza matematica e momenti

In questa Sezione ci occuperemo della definizione di media e dei momenti di variabili casuali assolutamente continue.

Definizione 5.4.4 Sia ξ una variabile casuale assolutamente continua di densità $f_\xi(x)$. Si dice che ξ ha speranza matematica (o media, o valore medio) finita se e solo se:

$$\int_{-\infty}^{+\infty} |x| f_\xi(x) dx < +\infty.$$

Se ξ ha speranza matematica finita, si chiama speranza matematica (o media, o valore medio) di ξ la quantità:

$$E(\xi) = \int_{-\infty}^{+\infty} x f_\xi(x) dx.$$

Anche per le variabili casuali assolutamente continue si possono dimostrare risultati del tutto analoghi ai Teoremi 5.3.3, 5.3.4 e 5.3.5 già enunciati nella Sezione 5.3 con riferimento alle variabili casuali discrete. Naturalmente, nel Teorema 5.3.3 la sommatoria viene sostituita dall'integrale della densità di probabilità, ottenendo:

$$E[\varphi(\xi)] = \int_{-\infty}^{+\infty} \varphi(x) f_\xi(x) dx.$$

Inoltre, la media è un operatore lineare e la media del prodotto di due variabili aleatorie assolutamente continue coincide con il prodotto delle rispettive medie solo se esse sono indipendenti.

Esempio 5.4.6

Si consideri una variabile casuale continua ξ uniforme sull'intervallo $[a, b]$ ($a, b \in \mathbb{R}$ con $a < b$). Come discusso nell'Esempio 5.4.1, ciò significa che $f_\xi(x)$ è costante e pari a k nell'intervallo $[a, b]$ e nulla altrove. Dalla condizione di normalizzazione:

$$1 = \int_{-\infty}^{+\infty} f_\xi(x) dx = \int_a^b k dx = k(b - a),$$

si ricava $f_\xi(x) = 1/(b - a)$. Il valore medio di ξ vale quindi:

$$E(\xi) = \int_{-\infty}^{+\infty} x f_\xi(x) dx = \int_a^b \frac{x}{b - a} dx = \frac{b^2 - a^2}{2(b - a)} = \frac{b + a}{2}.$$



Esempio 5.4.7

Si consideri una variabile casuale gaussiana ξ di densità $f_\xi(x) = N(0,1)$. Il suo valore medio è:

$$E(\xi) = \int_{-\infty}^{+\infty} x f_\xi(x) dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x e^{-x^2/2} dx = -\frac{1}{\sqrt{2\pi}} e^{-x^2/2} \Big|_{-\infty}^{+\infty} = 0.$$

Nel caso generico di una variabile gaussiana $\eta = \sigma\xi + \mu$ di parametri μ ed σ , per la linearità della media si ha:

$$E(\eta) = \sigma E(\xi) + \mu = \mu.$$



Anche i concetti momento e momento centrale di ordine k possono essere immediatamente estesi al caso delle variabili casuali assolutamente continue, ottenendo:

$$E(\xi^k) = \int_{-\infty}^{+\infty} x^k f_\xi(x) dx, \quad E[(\xi - E(\xi))^k] = \int_{-\infty}^{+\infty} [x - E(\xi)]^k f_\xi(x) dx.$$

In particolare, la varianza sarà il momento centrale di ordine 2:

$$\text{var}(\xi) = E[(\xi - E(\xi))^2] = E(\xi^2) - E(\xi)^2,$$

che gode delle proprietà ($a \in \mathbb{R}$):

$$\text{var}(a\xi) = a^2 \text{var}(\xi), \quad \text{var}(\xi + a) = \text{var}(\xi);$$

la grandezza $\sqrt{\text{var}(\xi)}$ viene detta *deviazione standard* della variabile aleatoria ξ .

Date due variabili casuali continue ξ e η , la loro covarianza è:

$$\text{cov}(\xi, \eta) = E[(\xi - E(\xi))(\eta - E(\eta))] = E(\xi\eta) - E(\xi)E(\eta),$$

che si annulla se le due variabili aleatorie sono indipendenti.

Esempio 5.4.8

Si consideri una variabile casuale continua ξ uniforme sull'intervallo $[a,b]$, quindi caratterizzata dalla densità $f_\xi(x) = 1/(b-a)$. La varianza di ξ vale:

$$\text{var}(\xi) = E(\xi^2) - E(\xi)^2 = \int_{-\infty}^{+\infty} x^2 f_\xi(x) dx - \frac{(b+a)^2}{4},$$

dove:

$$\int_{-\infty}^{+\infty} x^2 f_{\xi}(x) dx = \int_a^b x^2 f_{\xi}(x) dx = \frac{b^3 - a^3}{3(b-a)} = \frac{b^2 + ab + a^2}{3}.$$

Infine:

$$\text{var}(\xi) = \frac{b^2 + ab + a^2}{3} - \frac{(b+a)^2}{4} = \frac{(b-a)^2}{12}.$$

◆

Esempio 5.4.9

Si consideri una variabile casuale gaussiana ξ di densità $f_{\xi}(x) = N(0,1)$. La sua varianza è data da:

$$\text{var}(\xi) = E(\xi^2) - E(\xi)^2 = \int_{-\infty}^{+\infty} x^2 f_{\xi}(x) dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^2 e^{-x^2/2} dx.$$

Integrando per parti, si ha:

$$\int x^2 e^{-x^2/2} dx = -x e^{-x^2/2} + \int x e^{-x^2/2} dx,$$

per cui:

$$\text{var}(\xi) = -\frac{1}{\sqrt{2\pi}} x e^{-x^2/2} \Big|_{-\infty}^{+\infty} + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-x^2/2} dx = 1.$$

Nel caso generico di una variabile gaussiana $\eta = \sigma\xi + \mu$ di parametri μ ed σ , per le proprietà della varianza si ha:

$$\text{var}(\eta) = \text{var}(\sigma\xi + \mu) = \text{var}(\sigma\xi) = \sigma^2 \text{var}(\xi) = \sigma^2.$$

In definitiva, i parametri μ e σ di una variabile casuale gaussiana corrispondono al valore medio e alla deviazione standard della variabile stessa. ◆