

Appunti completi del corso di **ANALISI 2**

Equazioni Differenziali
Cenni sugli spazi vettoriali
Calcolo Differenziale

Ultima modifica: 1 marzo 2026

Indice

1 Equazioni Differenziali	4
1.1 Introduzione alle Equazioni Differenziali	4
1.2 EDO lineare	6
1.3 EDO lineare del primo ordine	7
1.3.1 Struttura della soluzione generale di una EDO lineare del primo ordine	7
1.3.2 Struttura dell'integrale generale di una EDO lineare del primo ordine omogenea associata	9
1.3.3 Struttura di una curva integrale di una EDO lineare del primo ordine completa	10
1.3.4 Problema di Cauchy per una EDO lineare del primo ordine	11
1.3.5 Intervallo massimale	12
1.4 EDO a variabili separabili	12
1.4.1 Le soluzioni singolari	14
1.5 EDO del secondo ordine	14
1.5.1 La struttura dello spazio delle soluzioni della EDO lineare di secondo ordine omogenea associata	14
1.5.2 Struttura dell'integrale generale della EDO lineare del secondo ordine omogenea associata	16
1.5.3 Struttura di una curva integrale di una EDO lineare del secondo ordine completa a coefficienti costanti	24
1.5.4 Problema di Cauchy per EDO lineare del secondo ordine a coefficienti costanti	27
2 Cenni di algebra lineare sugli spazi vettoriali	28
2.1 Il prodotto scalare e la norma	28
2.1.1 Disuguaglianze fondamentali indotte da una norma	29
2.2 Elementi di topologia indotti da una norma per uno spazio metrico	31
2.2.1 Successione nello spazio metrico	33
2.2.2 Intorno sferico	34
2.3 Classificazione topologica degli insiemi	35
2.3.1 I punti e gli insiemi	35
3 Funzioni di più variabili	38
3.1 Definizioni e grafici	38
3.1.1 Grafico di funzione e sezioni (trasversali e orizzontali)	38
3.1.2 Curve	39
3.1.3 Accenno al teorema degli zeri per funzioni a più variabili	40
3.2 Limiti e continuità	41
3.2.1 I limiti in più variabili e le loro proprietà	41
3.2.2 Continuità	43
3.2.3 Coordinate Polari	45
3.2.4 Il teorema del confronto per dimostrare l'esistenza di un limite	46

4 Calcolo Differenziale	48
4.1 Derivata	48
4.1.1 Derivata parziale	48
4.1.2 Gradiente	49
4.1.3 Derivata direzionale	50
4.2 Differenziabilità	51
4.2.1 Il concetto di funzione differenziabile e l'esistenza del piano tangente . . .	53
4.2.2 Continuità di una funzione differenziabile e il teorema del differenziale totale	54
4.2.3 La regola della catena	58
4.3 Derivate successive	60
4.3.1 Matrice Hessiana e il teorema di Schwarz	60
4.4 La formula di Taylor in funzioni multivariate	63
4.4.1 La formula di Taylor secondo il resto di Lagrange e di Peano	64
4.5 Studio di massimi e minimi per le funzioni a più variabili	66
4.6 Ottimizzazione libera	66
4.6.1 Test dell'hessiana	67
4.6.2 Forme Quadratiche	68
4.6.3 Caso dubbio: il test dell'hessiana non funziona	70
4.7 Ottimizzazione vincolata	71
4.7.1 Parametrizzazione	71
4.7.2 Moltiplicatori di Lagrange	71
4.7.3 Esempio di ottimizzazione vincolata con Sostituzione e Lagrange.	72
4.8 Integrazione	73
4.8.1 Integrale doppio e concetto di continuità	74
4.8.2 Calcolo Integrale sui rettangoli	74
4.8.3 Integrali doppi su domini generali e la misurabilità secondo Peano-Jordan	75
4.8.4 Proprietà degl integrali doppi	76
4.8.5 Calcolo Integrale su domini semplici	76
4.8.6 Domini simmetrici e la jacobiana	78

Equazioni Differenziali

1.1 Introduzione alle Equazioni Differenziali

Quando si definisce un'equazione differenziale, l'incognita è $y = y(x)$ che dipende dalla variabile $x \in I : [a, b] \subseteq \mathbb{R}$.

Definizione 1.1.1 (Equazione Differenziale). *Un'equazione la cui incognita è una funzione del tipo $y = y(x)$ e in cui sono presenti una o più derivate della stessa funzione incognita.*

Le equazioni differenziali si dividono in due categorie:

- **EDO**¹ (ordinarie), un'equazione differenziale su una funzione di una sola variabile indipendente;
- **EDP** (parziali), un'equazione differenziale su una funzione di più variabile indipendenti.

Nota. L'ordine di una EDO è l'ordine massimo di derivazione che compare nell'equazione differenziale.

Definizione 1.1.2 (EDO di ordine n). *Un'equazione differenziale ordinaria di ordine n si può scrivere in due forme:*

- *Implicita*

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

- $y(x) : I \subseteq \mathbb{R} \rightarrow \mathbb{R}$ è una funzione reale di variabile reale derivabile n volte in un intervallo I ;
- F è una funzione tale che $F : U \subseteq \mathbb{R}^{n+2} \rightarrow \mathbb{R}$, con U aperto, che prende in input $n+2$ numeri reali $(x, y(x), y'(x), y''(x), \dots, y^n(x))$ e produce in output un singolo numero reale.

- *Esplicita o Forma normale:*

$$y^n(x) = F(x, y(x), y'(x), y''(x), \dots, y^{n-1}(x))$$

dove $y(x) : I \subseteq \mathbb{R} \rightarrow \mathbb{R}$ e y^n è la derivata di ordine più alto.

Definizione 1.1.3 (Soluzione di una EDO). *Sia $F(x, y(x), y'(x), y''(x), \dots, y^n(x)) = 0$ con $x \in I \subseteq \mathbb{R}$ una EDO di ordine n su I .*

La sua soluzione è una funzione $\phi = \phi(x)$ definita su (almeno) $I = [a, b] \subseteq \mathbb{R}$ e $\phi \in C^n(I)$ per cui vale

$$F(x, \phi(x), \phi'(x), \phi''(x), \dots, \phi^n(x)) = 0$$

Definizione 1.1.4 (Soluzione generale e integrale generale di una EDO). *La soluzione generale di una EDO è una famiglia di funzioni che soddisfano l'equazione e che dipendono da una o più costanti arbitrarie (dette costanti di integrazione). Ogni scelta particolare delle costanti determina una soluzione particolare.*

¹Il corso approfondisce solo le EDO.

L'integrale generale di una EDO su un intervallo I è l'insieme di tutte le soluzioni dell'equazione definite su I .

Quando risolvo una EDO cerco tutte le possibili funzioni $y(x)$ che soddisfano l'uguaglianza tra la funzione incognita e le sue derivate.

Per formulare in modo più rigoroso il concetto, possiamo ricollegarci alla definizione di primitiva studiata in Analisi 1.

Sia $f : (a, b) \rightarrow \mathbb{R}$ una funzione assegnata. Si dice che una funzione $y : (a, b) \rightarrow \mathbb{R}$ è una *primitiva* di f in (a, b) se:

$$y'(x) = f(x), \quad \forall x \in (a, b)$$

L'uguaglianza $y'(x) = f(x)$ può essere interpretata come una EDO del primo ordine. In questa prospettiva, una primitiva di f non è altro che una soluzione della EDO

$$y' = f(x)$$

Possiamo quindi reinterpretare la definizione:

1. Si parte da una funzione assegnata $f(x)$;
2. Si cerca una funzione incognita $y(x)$;
3. Si verifica la condizione $y'(x) = f(x)$.

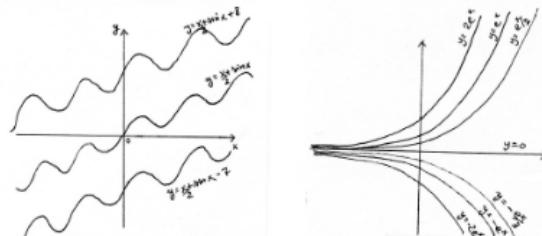
Se la derivata della funzione incognita candidata coincide identicamente con $f(x)$ su tutto l'intervallo, allora $y(x)$ è una soluzione dell'EDO, ossia una primitiva di f .

Dal punto di vista strutturale, l'insieme di tutte le primitive di f è descritto dall'integrale generale:

$$y(x, c) = \int f(x) dx + c, \quad c \in \mathbb{R}.$$

Questa espressione rappresenta una famiglia di funzioni dipendente da una costante arbitraria reale c . Abbiamo scritto che ogni valore della costante c individua una specifica soluzione della EDO.

Dal punto di vista geometrico, l'integrale generale rappresenta una famiglia di curve (integrali) nel piano (x, y) tali che, in ogni punto x , la loro pendenza è data dal valore $f(x)$.



Alcune soluzioni delle EDO del primo ordine $y'(x) = \frac{1}{2} + \cos(x)$ e $y'(x) = y$

In conclusione, l'integrale generale di una EDO è l'insieme di tutte le sue soluzioni, espresso in forma parametrica come $y(x, c)$, dove la variabile indipendente è x e il parametro reale c distingue le diverse soluzioni appartenenti alla stessa famiglia.

1.2 EDO lineare

Una EDO è non lineare se

- $y(x)$ e le sue derivate compaiono dentro funzioni "complesse", per esempio potenze y^2 , seno $\sin(y)$, coseno $\cos(y)$, ...;
- $y(x)$ e le sue derivate sono moltiplicate tra loro, per esempio $y'(x) \cdot y(x) = 0$;
- i coefficienti a dipendono dalla funzione $y(x)$, per esempio $y'(x)y(x) + y(x) = 0$.

Nota. Per parlare di linearità di una EDO bisogna interpretarla come un'operazione che agisce fra funzioni.

Definizione 1.2.1 (Linearità di una EDO). Dato $I \subseteq \mathbb{R}$ e consideriamo lo spazio $C^n(I)$ delle funzioni n -volte derivabili su I .

Un operatore differenziale lineare di ordine n è un'applicazione lineare $L : C^n(I) \rightarrow C^0(I)$.

Una EDO in forma normale $y^n = F(x, y, y', y'', \dots, y^{n-1})$ è definita lineare perché può essere scritta nella forma

$$L(\phi(x)) := a_n \phi^n(x) + a_{n-1} \phi^{n-1}(x) + \dots + a_1 \phi'(x) + a_0 \phi(x)$$

dove L associa ad ogni funzione $\phi \in C^n(I)$ la funzione $L(\phi) \in C^0(I)$, con $a_n, a_{n-1}, \dots, a_1, a_0$ funzioni continue su I .

L'applicazione lineare L è lineare perché soddisfa due proprietà fondamentali:

- L'additività

$$L(\phi_1 + \phi_2) = L\phi_1 + L\phi_2, \quad \forall \phi_1, \phi_2 \in C^n(I)$$

- L'omogeneità

$$L(\alpha \phi) = \alpha L\phi, \quad \forall \phi \in C^n(I), \forall \alpha \in \mathbb{R}$$

Insieme si riasusmono in una sola proprietà di linearità:

$$L(\alpha \phi_1 + \beta \phi_2) = \alpha L(\phi_1) + \beta L(\phi_2), \quad \forall \alpha, \beta \in \mathbb{R}, \forall \phi_1, \phi_2 \in C^n(I)$$

In parole povere, la linearità significa che

- non si creano relazioni strane tra le funzioni e non compaiono prodotti come $y(x) \cdot y'(x)$ o come potenze y^2 o funzioni non lineari come $\sin(y)$, $\cos(y)$, ...;
- ogni termine dipende da y o dalle sue derivate in modo separato;
- l'operatore L le elabora separatamente e ne somma i risultati.

Le funzioni sono trattate in modo indipendente e proporzionale (per esempio, se raddoppi la funzione in ingresso, la funzione in uscita raddoppia ugualmente).

Questa proprietà è fondamentale perché implica il principio di sovrapposizione per le equazioni lineari omogenee, ogni combinazione lineare di soluzioni è ancora una soluzione.

Definizione 1.2.2 (EDO lineare Completa). Sia il termine noto $f(x) \neq 0$, una EDO lineare di ordine n può essere scritta in due forme

- Implicita: $a_n y^n(x) + a_{n-1} y^{n-1}(x) + \dots + a_1 y'(x) + a_0 y(x) = f(x)$
- Esplicita o forma normale: $y^n(x) + a_{n-1} y^{n-1}(x) + \dots + a_1 y'(x) + a_0 y(x) = f(x)$

Definizione 1.2.3 (EDO lineare Omogenea Associata ad una EDO completa). Sia il termine noto uguale a zero, si scrive nella forma:

$$a_n y^n(x) + a_{n-1} y^{n-1}(x) + \dots + a_1 y'(x) + a_0 y(x) = 0$$

1.3 EDO lineare del primo ordine

Definizione 1.3.1 (EDO lineare del primo ordine). Sia $x \in [a, b]$ e siano $a(x)$ e $f(x)$ funzioni continue in $[a, b]$, posso definire la EDO lineare del primo ordine in due forme

- *Forma normale:* $y'(x) + a(x)y(x) = f(x)$
- *Omogena associata:* $y'(x) + a(x)y(x) = 0$

La continuità di $a(x)$ e $f(x)$ è fondamentale poiché garantisce l'esistenza e l'unicità della soluzione del problema di Cauchy associato.

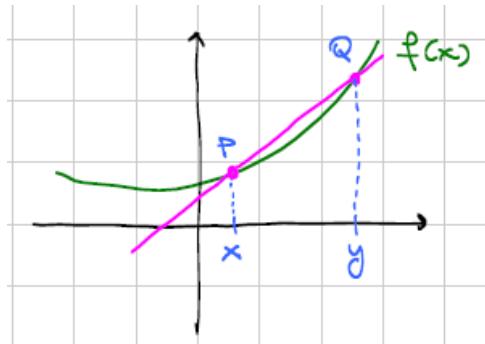
Consideriamo il problema di Cauchy:

$$\begin{cases} y' = f(x, y) \\ y(x_0) = y_0 \end{cases}$$

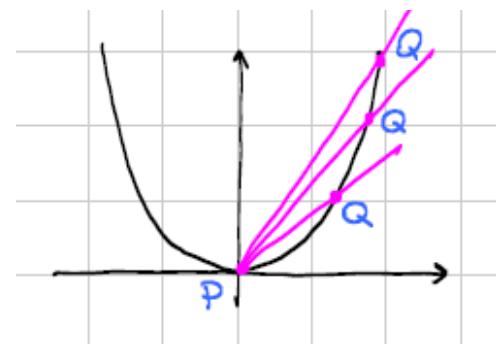
La soluzione esiste ed è unica in un intorno del punto (x_0, y_0) se

- **Esistenza (Continuità):** la funzione $f(x, y)$ è continua in un intorno del punto (x_0, y_0) ;
- **Unicità (Condizione di Lipschitz²):** la funzione f è lipschitziana rispetto a y se la derivata parziale $\frac{\partial f}{\partial y}$ esiste ed è continua nell'intorno considerato.

Questa condizione garantisce che le soluzioni del problema di Cauchy partendo da uno stesso punto iniziale non possano separarsi, assicurando l'unicità della soluzione. Senza la condizione di Lipschitz possono esistere più soluzioni distinte con la stessa condizione iniziale;



Funzione Lipschitziana



Funzione non Lipschitziana

Nota. La continuità di $a(x)$ e $f(x)$ assicura inoltre la possibilità di applicare la formula risolutiva che richiede il calcolo di integrali di tali funzioni.

1.3.1 Struttura della soluzione generale di una EDO lineare del primo ordine

Teorema 1.3.1 (Rappresentazione della soluzione generale di una EDO lineare completa). La soluzione generale di una EDO lineare del primo ordine definita su un intervallo $I = [a, b]$ del tipo $y'(x) + a(x)y(x) = f(x)$ è dato dalla somma dell'integrale generale dell'omogenea associata ($y'(x) + a(x)y(x) = 0$) ed una curva integrale (una soluzione particolare) della EDO lineare non omogenea in forma normale ($y'(x) + a(x)y(x) = f(x)$):

$$y(x) = y_{om}(x) + \bar{y}(x)$$

dove $x \in I$ e $y(x)$ è una soluzione qualsiasi della EDO lineare di ordine n , $y_{om}(x)$ è una soluzione qualsiasi della EDO lineare omogenea associata di ordine n e $\bar{y}(x)$ è una soluzione particolare della EDO lineare di ordine n

²La variazione di f rispetto a y è controllata. Geometricamente, il grafico non può avere pendenze infinite né oscillazioni troppo brusche.

DIMOSTRAZIONE per $n = 1$ (è valida anche per n di ordine superiore)

Parte 1 Mostriamo che $y_{\text{om}}(x) = y(x) - \bar{y}(x)$ è soluzione dell'omogenea.

$$\begin{aligned} y'(x) + a(x)y(x) &= f(x) \quad (\text{soluzione qualsiasi della completa}) \\ \bar{y}'(x) + a(x)\bar{y}(x) &= f(x) \quad (\text{soluzione particolare della completa}) \end{aligned}$$

Sottraiamo la seconda equazione dalla prima, membro a membro:

$$\begin{aligned} [y'(x) + a(x)y(x)] - [\bar{y}'(x) + a(x)\bar{y}(x)] &= f(x) - f(x) \\ y'(x) + a(x)y(x) - \bar{y}'(x) - a(x)\bar{y}(x) &= 0 \\ y'(x) - \bar{y}'(x) + a(x)y(x) - a(x)\bar{y}(x) &= 0 \quad (\text{ri-ordino}) \\ [y'(x) - \bar{y}'(x)] + a(x)[y(x) - \bar{y}(x)] &= 0 \quad (\text{raccolgo } a(x)) \end{aligned}$$

Usando la linearità della derivata: $y' - \bar{y}' = (y - \bar{y})'$, otteniamo:

$$[y(x) - \bar{y}(x)]' + a(x)[y(x) - \bar{y}(x)] = 0$$

Ponendo $y_{\text{om}}(x) = y(x) - \bar{y}(x)$:

$$y'_{\text{om}}(x) + a(x)y_{\text{om}}(x) = 0$$

Quindi $y_{\text{om}}(x)$ è soluzione dell'omogenea perché abbiamo dimostrato che la differenza è soluzione dell'equazione omogenea associata.

Parte 2 Mostriamo che $y(x) = y_{\text{om}}(x) + \bar{y}(x)$ è soluzione della completa.

$$\begin{aligned} y'_{\text{om}}(x) + a(x)y_{\text{om}}(x) &= 0 \quad (\text{soluzione dell'omogenea}) \\ \bar{y}'(x) + a(x)\bar{y}(x) &= f(x) \quad (\text{soluzione particolare della completa}) \end{aligned}$$

Sommiamo le due equazioni membro a membro:

$$\begin{aligned} [y'_{\text{om}}(x) + a(x)y_{\text{om}}(x)] + [\bar{y}'(x) + a(x)\bar{y}(x)] &= 0 + f(x) \\ y'_{\text{om}}(x) + a(x)y_{\text{om}}(x) + \bar{y}'(x) + a(x)\bar{y}(x) &= f(x) \\ y'_{\text{om}}(x) + \bar{y}'(x) + a(x)y_{\text{om}}(x) + a(x)\bar{y}(x) &= f(x) \quad (\text{ri-ordino}) \\ [y'_{\text{om}}(x) + \bar{y}'(x)] + a(x)[y_{\text{om}}(x) + \bar{y}(x)] &= f(x) \quad (\text{raccolgo } a(x)) \end{aligned}$$

Usando la linearità della derivata $y'_{\text{om}} + \bar{y}' = (y_{\text{om}} + \bar{y})'$, otteniamo:

$$[y_{\text{om}}(x) + \bar{y}(x)]' + a(x)[y_{\text{om}}(x) + \bar{y}(x)] = f(x)$$

Ponendo $y(x) = y_{\text{om}}(x) + \bar{y}(x)$:

$$y'(x) + a(x)y(x) = f(x)$$

Quindi $y(x) = y_{\text{om}}(x) + \bar{y}(x)$ è soluzione della completa. □

La rappresentazione alternativa dell'integrale generale $y'(x) + a(x)y(x) = f(x)$ è

$$\int y(x) = \int y_{\text{om}}(x) + \int \bar{y}(x)$$

Nota. Per trovare tutte le soluzioni della EDO lineare è necessario sommare la soluzione generale dell'omogenea associata e la soluzione nota dell'equazione di partenza.

1.3.2 Struttura dell'integrale generale di una EDO lineare del primo ordine omogenea associata

Teorema 1.3.2 (L'integrale generale della EDO lineare del primo ordine omogenea associata). *Sia $y'(x) + a(x)y(x) = 0$ con $x \in [a, b]$ e $a(x)$ continua in $[a, b]$, l'integrale generale dell'omogenea associata prende la forma*

$$y(x) = y(x, c) = ce^{-A(x)} = ce^{-\int a(x)dx}, \quad c \in \mathbb{R}$$

con $A(x) = \int a(x)dx$ primitiva di $a(x)$ fissata una volta per tutte.

DIMOSTRAZIONE

Si moltiplica ambo i membri di $y'(x) + a(x)y(x) = 0$ per $e^{A(x)} = e^{\int a(x)dx}$ perché l'esponenziale ha la derivata proporzionale a se stesso e permette di trasformare l'omogenea associata in una derivata del prodotto, e quindi si ottiene

$$\underbrace{e^{A(x)}y'(x) + e^{A(x)}a(x)y(x)}_{[e^{A(x)}y(x)]'} = 0,$$

essendo la derivata di una costante uguale 0 ci dice che

$$e^{A(x)}y(x) = c$$

$$\frac{e^{A(x)}y(x)}{e^{A(x)}} = \frac{c}{e^{A(x)}}$$

$$y(x) = \frac{c}{e^{A(x)}}$$

$$y(x) = ce^{-A(x)}$$

allora (è ottenuto l'integrale generale dell'omogenea associata)

$$y(x) = ce^{-A(x)} = ce^{-\int a(x)dx}, \quad c \in \mathbb{R}.$$

Infatti

$$(e^{-A(x)})' = e^{-A(x)}[-a(x)],$$

dunque

$$(e^{-A(x)})' + e^{-A(x)}a(x) = 0.$$

□

Nota. L'insieme delle soluzioni della EDO lineare del primo ordine dell'omogenea associata costituisce uno spazio vettoriale³ di dimensione 1 e per trovare gli elementi di tale spazio:

- Basta trovare una soluzione particolare non nulla $y_{om}(x) = e^{-A(x)}$;
- Questa $y_{om}(x)$ funge da base dello spazio;
- Ogni altra soluzione è uno suo multiplo e si può scrivere come $y(x) = ce^{-A(x)}$

³Questo perché soddisfa le proprietà fondamentali della linearità: la somma di due soluzioni è ancora una soluzione e il prodotto di una soluzione per una costante è una soluzione

1.3.3 Struttura di una curva integrale di una EDO lineare del primo ordine completa

Dopo aver strutturato la soluzione generale della EDO lineare del primo ordine omogenea associata, il passo successivo è quello di trovare una curva integrale⁴ $\bar{y}(x)$ della stessa EDO lineare completa.

Essa può essere strutturata "ad occhio", analizzando il termine noto, o con il metodo di variazione di una costante.

Teorema 1.3.3 (Una curva integrale della EDO lineare del primo ordine completa). *Sia $y'(x) + a(x)y(x) = f(x)$ la struttura di una soluzione particolare della EDO lineare completa è definita dal metodo di variazione di una costante nella forma*

$$\bar{y}(x) = e^{-A(x)} \int f(x)e^{A(x)} dx$$

DIMOSTRAZIONE

Consideriamo la soluzione nella forma:

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

dove $e^{-A(x)}$ è la soluzione dell'omogenea associata e $c(x) \in C^1(I)$ ed è una funzione perché non mi basta più una costante c fissa per strutturare la soluzione particolare.

Essendo $\bar{y}(x)$ una soluzione particolare della EDO completa del primo ordine, è possibile trasformare la

$$\bar{y}'(x) + a(x)\bar{y}(x) = f(x), \quad x \in [a, b].$$

Sostituendo al suo interno $\bar{y}(x) = c(x)e^{-A(x)}$, ottengo la EDO nella forma

$$(c(x)e^{-A(x)})' + a(x)c(x)e^{-A(x)}$$

Quindi, si deriva $(c(x)e^{-A(x)})'$ e la forma completa diventa

$$c'(x)e^{-A(x)} - c(x)a(x)e^{-A(x)} + a(x)c(x)e^{-A(x)} = f(x),$$

Con un passaggio algebrico si trasforma

$$c'(x)e^{-A(x)} - \underline{c(x)a(x)e^{-A(x)}} + \underline{a(x)c(x)e^{-A(x)}} = f(x),$$

ovvero

$$c'(x)e^{-A(x)} = f(x),$$

Dividendo tutto per $e^{-A(x)}$ e poi spostando $e^{-A(x)}$ al numeratore, ottengo la forma definitiva

$$c'(x) = e^{A(x)}f(x).$$

Integrando $c'(x)$, infine, si ottiene:

$$c(x) = \int e^{A(x)}f(x) dx.$$

Pertanto, sostituendo in $\bar{y}(x) = c(x)e^{-A(x)}$, si ottiene la definizione dell'integrale generale particolare $\bar{y}(x)$ della EDO completa non omogenea nella forma

$$\bar{y}(x) = e^{-A(x)} \int e^{A(x)}f(x) dx.$$

□

Per concludere, per calcolare l'integrale generale della EDO non omogenea lineare si usa

$$y(x) = ce^{-A(x)} + e^{-A(x)} \int f(x)e^{A(x)} dx$$

Nota. $A(x)$ è una primitiva di $a(x)$ che viene scelta una e una sola volta e non occorre aggiungere una costante arbitraria in quanto, nella formula finale, è già inclusa la costante c ; non occorre aggiungere nemmeno una costante additiva nel calcolo dell'integrale particolare. Tutto ciò è dimostrabile ripetendo le dimostrazione con l'aggiunta di un k .

⁴Per curva particolare si intende una soluzione particolare che soddisfa la EDO completa.

1.3.4 Problema di Cauchy per una EDO lineare del primo ordine

$$\begin{cases} y'(x) + a(x)y(x) = f(x) \\ y(x_0) = y_0 \end{cases}$$

Teorema 1.3.4 (Il problema di Cauchy per EDO lineari del primo ordine). *Sia $y'(x) + a(x)y(x) = f(x)$ con $x \in [a, b]$ e $a(x), f(x)$ continue in $[a, b]$, allora il problema di Cauchy associato ammette una soluzione unica definita su tutto l'intervallo $[a, b]$ ed è data da*

$$y(x) = y_0 e^{-\int_{x_0}^x a(t) dt} + e^{-\int_{x_0}^x a(t) dt} \int_{x_0}^x e^{\int_{x_0}^s a(t) ds} f(s) ds$$

DIMOSTRAZIONE

Premettiamo che il problema di Cauchy per EDO lineari ci permette di determinare la costante c che compare nell'integrale generale dell'omogenea associata, imponendo una condizione iniziale $y(x_0) = y_0$ con $x_0 \in [a, b]$.

Definiamo la primitiva $A(x)$ attraverso l'integrale definito con estremo inferiore x_0 :

$$A(x) = \int_{x_0}^x a(t) dt$$

Per il teorema fondamentale del calcolo integrale, questa scelta implica che:

$$A(x_0) = \int_{x_0}^{x_0} a(t) dt = 0$$

Sostituendo tale primitiva nella soluzione generale $y(x) = ce^{-A(x)}$, otteniamo per $x = x_0$:

$$y(x_0) = ce^{-A(x_0)} = ce^0 = c$$

Quindi, moltiplichiamo ambedue i membri della EDO per il fattore integrante $e^{A(x)}$:

$$e^{A(x)} y'(x) + a(x)e^{A(x)} y(x) = e^{A(x)} f(x)$$

Osserviamo che il membro di sinistra è lo sviluppo della derivata del prodotto tra la funzione incognita e il fattore integrante:

$$\frac{d}{dx} \left[e^{A(x)} y(x) \right] = e^{A(x)} f(x)$$

Integrando entrambi i membri rispetto alla variabile di integrazione s nell'intervallo $[x_0, x]$ otteniamo:

$$\int_{x_0}^x e^{A(s)} y(s) ds = \int_{x_0}^x e^{A(s)} f(s) ds$$

Applicando il Teorema Fondamentale del Calcolo Integrale al membro di sinistra:

$$\begin{aligned} \left[e^{A(s)} y(s) \right]_{x_0}^x &= \int_{x_0}^x e^{A(s)} f(s) ds \\ e^{A(x)} y(x) - e^{A(x_0)} y(x_0) &= \int_{x_0}^x e^{A(s)} f(s) ds \end{aligned}$$

Poiché abbiamo definito $A(x_0) = 0$, ne consegue che $e^{A(x_0)} = e^0 = 1$. Sostituendo la condizione iniziale $y(x_0) = y_0$:

$$e^{A(x)} y(x) - y_0 = \int_{x_0}^x f(s) e^{A(s)} ds$$

Isolando $y(x)$ e moltiplicando tutto per $e^{-A(x)}$, si ottiene la formula risolutiva esplicita:

$$y(x) = y_0 e^{-\int_{x_0}^x a(t) dt} + e^{-\int_{x_0}^x a(t) dt} \int_{x_0}^x e^{-\int_{x_0}^s a(t) ds} f(s) ds = e^{-A(x)} \left[y_0 + \int_{x_0}^x f(s) e^{A(s)} ds \right]$$

□

1.3.5 Intervallo massimale

- **Soluzione in piccolo (locale):** La soluzione esiste ed è definita solo in un intorno del punto iniziale $x_0 \in I$, ovvero l'intervallo in cui è definita l'EDO;
- **Soluzione in grande (globale):** La soluzione esiste ed è definita su tutto l'intervallo di definizione delle funzioni coefficienti, che può essere tutto \mathbb{R} o un intervallo più ampio.

L'intervallo massimale di esistenza di una soluzione è il più grande intervallo contenente il punto iniziale x_0 nel quale la soluzione è definita e soddisfa la EDO.

Può essere limitato o illimitato:

- **Può essere tutto \mathbb{R}** se le funzioni coefficienti $a(x)$ e $f(x)$ sono definite e continue su tutto \mathbb{R} e la soluzione non presenta singolarità.
- **Può essere limitato** se:
 - Le funzioni coefficienti non sono definite su tutto \mathbb{R} (ad esempio: se $a(x) = \frac{1}{\sqrt{x}}$, allora $I = (0, +\infty)$)
 - La soluzione presenta singolarità o comportamenti asintotici che impediscono il prolungamento

Da cosa dipende:

1. Dall'**intervallo di definizione** delle funzioni $a(x)$ e $f(x)$
2. Dal **comportamento della soluzione** stessa (presenza di singolarità, esplosioni in tempo finito)
3. Dalle **condizioni iniziali** (x_0, y_0) che possono influenzare quando e dove la soluzione cessa di esistere

L'intervallo massimale è determinato dalla forma delle funzioni coinvolte nell'EDO e dal punto iniziale scelto. Non è detto che le soluzioni espresse dall'integrale generale dell'equazione siano tutte definite sullo stesso insieme al variare dei parametri o delle condizioni iniziali.

- **Fisico:** Il problema di Cauchy rappresenta un'evoluzione rispetto al tempo (o rispetto alla variabile indipendente x), che è una variabile continua. Quindi se è trovato un tempo in cui il sistema perde significato non ha senso considerare il prima o il dopo. Allora viene considerato l'intervallo massimale contenente l'istante iniziale in cui il sistema ha significato.
- **Matematico:** Non è possibile considerare una soluzione $y(x)$ al PdC su intervalli disgiunti. Per questo, la condizione iniziale determina l'esistenza della soluzione e l'unicità di essa nell'insieme che la contiene.

1.4 EDO a variabili separabili

Definizione 1.4.1 (EDO a variabili separabili del primo ordine). Una EDO a variabili separabili in forma normale si scrive nella forma

$$y'(x) = f(x)g(y(x))$$

dove $f : I \in \mathbb{R} \rightarrow \mathbb{R}$ e $g : J \in \mathbb{R} \rightarrow \mathbb{R}$ sono continue nei loro domini.

Nota (Una EDO lineare omogenea del primo ordine è sempre EDO a variabili separabili). Una EDO lineare omogenea associata del primo ordine ha la forma:

$$y'(x) + a(x)y(x) = 0$$

Riscriviamola esplicitamente come:

$$\frac{dy}{dx} = -a(x)y(x)$$

Questa può essere separata dividendo per y (assumendo $y \neq 0$) e moltiplicando per dx :

$$\frac{dy}{y} = -a(x) dx$$

Ora le variabili sono separate: a sinistra compare solo y , a destra solo x . Integrando entrambi i membri:

$$\int \frac{dy}{y} = \int -a(x) dx \Rightarrow \ln |y| = -A(x) + c$$

dove $A(x) = \int a(x) dx$. Esponenziando:

$$y(x) = ce^{-A(x)}$$

Quindi, una EDO lineare omogenea del primo ordine è sempre riconducibile alla forma separabile.

Teorema 1.4.1 (Algoritmo per la separazione delle variabili). *Partendo da $y'(x) = f(x)g(y(x))$*

1. Determina eventuali $\bar{y} \in \mathbb{R}$ per cui $g(\bar{y}) = 0$ dove $y(x) = \bar{y}$ è la soluzione del PdC. Queste sono le soluzioni singolari della EDO;
2. Si considera $y \neq \bar{y}$ si divide tutto per $g(y(x))$ separando, quindi, la EDO (dove $g(y(x)) \neq 0$ dato che $y \neq \bar{y}$)

$$\frac{y'(x)}{g(y(x))} = \frac{f(x)\cancel{g(y(x))}}{\cancel{g(y(x))}}$$

3. Si integra ambo ai lati per x per ottenere le soluzioni

$$\int \frac{y'(x)}{g(y(x))} dx = \int f(x) dx$$

4. Si usa la regola della sostituzione ponendo $y = y(x)$ e $dy = y'(x)dx$

$$\int \frac{1}{g(y)} dy = \int f(x) dx$$

5. Sia $G(x)$ la primitiva di $\frac{1}{g(y)}$ e sia $F(x)$ di $f(x)$

$$G(x) = F(x) + c$$

6. Per concludere, si esplicita $y(x)$ è la soluzione del PdC, con G invertibile si ha

$$y(x) = G^{-1}(x)[F(x) + c]$$

□

1.4.1 Le soluzioni singolari

Le soluzioni singolari sono soluzioni della EDO che non sono ottenibili dall'integrale generale per nessun valore della costante arbitraria c .

Si individuano attraverso le equazioni a variabili separabili della forma:

$$\frac{dy}{dx} = f(x)g(y(x))$$

Le soluzioni singolari si ottengono dagli zeri della funzione $g(y(x))$. Se $g(y_0) = 0$, allora $y(x) = y_0$ (funzione costante) è una soluzione dell'equazione.

Durante la separazione delle variabili, quando scriviamo:

$$\frac{dy}{g(y)} = f(x) dx$$

dividiamo per $g(y)$, e questa operazione è valida solo se $g(y) \neq 0$. Pertanto, le soluzioni corrispondenti a $g(y) = 0$ vengono perse nel processo di separazione e devono essere verificate separatamente.

Le soluzioni singolari sono importanti perché la loro presenza può influenzare l'unicità delle soluzioni di problemi di Cauchy, infatti rappresentano soluzioni effettive dell'EDO.

1.5 EDO del secondo ordine

Definizione 1.5.1 (EDO lineare completa del secondo ordine). Una EDO lineare del secondo ordine viene rappresentata nella forma

$$a_2(x)y''(x) + a_1(x)y'(x) + a_0(x)y(x) = f(x)$$

dove a_i (con $i = 0, 1, 2$) e $f(x)$ sono costanti (ovvero appartengono a $C^0(I)$)

Definizione 1.5.2. Basandoci sulla forma della EDO lineare del secondo ordine, ponendo $a_2(x) = 1$, possiamo scrivere l'EDO in forma normale e la sua omogenea associata

1.

$$y''(x) + a_1(x)y'(x) + a_0(x)y(x) = f(x)$$

2.

$$y''(x) + a_1(x)y'(x) + a_0(x)y(x) = 0$$

Per calcolare una EDO lineare del II ordine, come per una EDO del primo ordine, bisogna trovare l'integrale generale dell'omogenea associata e sommarlo all'integrale particolare della non omogenea, ottenendo l'integrale della completa:

$$\int \text{generale} = \int \text{omogenea} + \int \text{particolare}$$

Nota. Lo spazio delle soluzioni di una EDO di ordine $n = 2$ è uno spazio vettoriale di dimensione $n = 2$ per trovare una soluzione qualsiasi dell'omogenea servono $n = 2$ soluzioni linearmente indipendenti.

1.5.1 La struttura dello spazio delle soluzioni della EDO lineare di secondo ordine omogenea associata

Spazio delle soluzioni

Se otteniamo come soluzioni su $I = [a, b]$ della omogenea associata $y''(x) + a_1(x)y'(x) + a_0(x)y(x) = 0$ le funzioni $y_1(x)$ e $y_2(x)$ possiamo dire che entrambe sono linearmente indipendenti e che ogni altra soluzione è una combinazione lineare delle due "funzioni-soluzioni".

Il motivo è la proprietà della linearità della EDO: dato un operatore $L : C^n(I) \rightarrow C^0(I)$ che associa ad ogni funzione $\phi \in C^n(I)$ una funzione continua $L(\phi(x)) := a_n\phi^n(x) + a_{n-1}\phi^{n-1}(x) + \dots + a_1\phi'(x) + a_0\phi(x)$. Segue che se ϕ_1 e ϕ_2 sono soluzioni, anche la loro combinazione lineare è soluzione, allora l'integrale generale della omogenea associata è rappresentabile come

$$y_{\text{om}}(x, c_1, c_2) = c_1 y_1(x) + c_2 y_2(x), \quad c_1, c_2 \in \mathbb{R}$$

La EDO lineare omogenea del secondo ordine può essere scritta in forma compatta tramite l'operatore

$$L(\phi(x)) := a\phi''(x) + b\phi'(x) + c\phi(x), \quad \phi \in C^2(I)$$

Tutto ciò significa che risolvere $ay''(x) + by'(x) + cy(x) = 0$ significa trovare la y tale $ay''(x) + by'(x) + cy(x) = L(y(x)) = 0$.

Wroskiano: indipendenza lineare

Definizione 1.5.3 (Matrice Wroskiana). Date due funzioni differenziabili $y_1(x)$ e $y_2(x)$, la matrice Wroskiana è

$$\begin{vmatrix} y_1(x) & y_2(x) \\ y'_1(x) & y'_2(x) \end{vmatrix}$$

Nota. Il determinante della matrice Wroskiana è uno "strumento" che permette di verificare se due funzioni sono linearmente indipendenti su un intervallo.

Definizione 1.5.4 (Wroskiano). Il determinante Wroskiano o Wronskiano è il determinante di una matrice formata da due funzioni e dalle loro derivate prime: $W[y_1, y_2] = \det(W(x)) =$

$$\begin{vmatrix} y_1(x) & y_2(x) \\ y'_1(x) & y'_2(x) \end{vmatrix}$$

$$= y_1(x)y'_2(x) - y'_1(x)y_2(x)$$

Se il Wronskiano è diverso da zero in almeno un punto dell'intervallo, allora le funzioni sono linearmente indipendenti su quell'intervallo. In altre parole:

$$W[y_1, y_2](x) \neq 0 \text{ per qualche } x \in I \iff y_1, y_2 \text{ sono linearmente indipendenti su } I$$

Sistema fondamentale di soluzioni

Definizione 1.5.5 (Sistema fondamentale di soluzioni). Un insieme $\{y_1(x), y_2(x)\}$ di soluzioni di una EDO lineare omogenea del secondo ordine si dice sistema fondamentale di soluzioni su un intervallo I se:

1. y_1 e y_2 sono linearmente indipendenti su I ;
2. ogni soluzione dell'equazione omogenea può essere espressa come combinazione lineare di y_1 e y_2 .

Il sistema fondamentale di soluzioni è importante perché fornisce una base completa per lo spazio delle soluzioni dell'equazione differenziale omogenea. In particolare:

- **Analogia con l'algebra lineare:** Così come ogni vettore di \mathbb{R}^n può essere rappresentato come combinazione lineare di una base (ad esempio, $(1, 2) = 1\underline{e}_1 + 2\underline{e}_2$ in \mathbb{R}^2), ogni soluzione di un'EDO lineare omogenea di ordine n può essere espressa come combinazione lineare di n soluzioni linearmente indipendenti;
- **Struttura di spazio vettoriale:** Lo spazio delle soluzioni di un'EDO lineare omogenea del secondo ordine ha dimensione 2. Pertanto, due soluzioni linearmente indipendenti formano una base di questo spazio;
- **Semplicità:** Per caratterizzare completamente lo spazio delle soluzioni, è sufficiente trovare sole due soluzioni linearmente indipendenti, anziché tutte le infinite soluzioni.

1.5.2 Struttura dell'integrale generale della EDO lineare del secondo ordine omogenea associata

La forma della omogenea associata è $ay''(x) + by'(x) + cy(x) = 0$

1. Se $b = c = 0$ (identicamente nulle) l'equazione diventa $ay'' = 0$, quindi $y'(x) = c_1$ integrando si ottiene $y_{om}(x) = z(x, c_1, c_2) = c_1 + c_2$ con $c_1, c_2 \in \mathbb{R}$
2. Se b e c non sono contemporaneamente nulle, per trovare l'integrale generale della omogenea viene associata alla EDO un'equazione algebrica $p(\lambda)$ del secondo ordine in \mathbb{C} della forma

$$a\lambda^2 + b\lambda + c = 0, \quad \lambda \in \mathbb{R}$$

Formalmente le soluzioni di $p(\lambda)$ sono date dalla formula risolutiva $\lambda_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$ che ci permettono di definire tre disinti casi:

1. **Caso $\Delta > 0$**

Soluzioni reali distinte: $\lambda_{1,2} \in \mathbb{R}$ e $\lambda_1 \neq \lambda_2$

$$y_1(x) = e^{\lambda_1 x} \text{ e } y_2(x) = e^{\lambda_2 x}$$

L'integrale generale dell'omogenea:

$$y_{om}(x) = c_1(x)y_1(x) + c_2(x)y_2(x) = c_1 e^{\lambda_1 x} + c_2 e^{\lambda_2 x}$$

2. **Caso $\Delta = 0$**

Soluzioni reali coincidenti: $\lambda_1 = \lambda_2 = \lambda = \frac{-b}{2a} \in \mathbb{R}$

L'integrale generale dell'omogenea:

$$y_{om}(x) = e^{\lambda x}(c_1 + c_2 x)$$

3. **Caso $\Delta < 0$**

Soluzioni complesse coniugate: $\lambda_{1,2} = \alpha \pm i\beta$ con $i = \sqrt{-1}$, $\alpha = -\frac{b}{2a}$ e $\beta = \frac{\sqrt{-\Delta}}{2a}$

L'integrale generale dell'omogenea:

$$y_{om}(x) = e^{\alpha x}(c_1 \cos \beta x + c_2 \sin \beta x)$$

Teorema 1.5.1 (λ è radice del polinomio caratteristico se e solo se $e^{\lambda x}$ è soluzione dell'omogenea associata). Sia $p(\lambda) = a\lambda^2 + b\lambda + c$ il polinomio caratteristico associato alla EDO lineare omogenea del secondo ordine a coefficienti costanti $ay''(x) + by'(x) + cy(x) = 0$.

Allora, $\lambda \in \mathbb{C}$ è radice di $p(\lambda)$ se e solo se la funzione $y(x) = e^{\lambda x}$ è soluzione dell'equazione omogenea.

DIMOSTRAZIONE

Si consideri l'operatore differenziale lineare del secondo ordine a coefficienti costanti:

$$L(y) = ay'' + by' + cy$$

Vogliamo dimostrare che una funzione esponenziale del tipo $y(x) = e^{\lambda x}$ è soluzione dell'equazione omogenea $L(y) = 0$ se e solo se λ è radice del polinomio caratteristico $p(\lambda)$.

Sostituiamo la funzione $y(x) = e^{\lambda x}$ nell'operatore L . Calcoliamo innanzitutto le derivate della funzione:

$$y(x) = e^{\lambda x}, \quad y'(x) = \lambda e^{\lambda x}, \quad y''(x) = \lambda^2 e^{\lambda x}$$

Sostituendo nell'espressione dell'operatore $L(e^{\lambda x})$ otteniamo:

$$L(e^{\lambda x}) = a(\lambda^2 e^{\lambda x}) + b(\lambda e^{\lambda x}) + c(e^{\lambda x})$$

Possiamo ora raccogliere il termine comune $e^{\lambda x}$:

$$L(e^{\lambda x}) = e^{\lambda x}(a\lambda^2 + b\lambda + c)$$

Definiamo il polinomio caratteristico come $p(\lambda) = a\lambda^2 + b\lambda + c$. L'espressione diventa:

$$L(e^{\lambda x}) = e^{\lambda x} \cdot p(\lambda)$$

Poiché la funzione esponenziale $e^{\lambda x}$ non è mai nulla per alcun valore reale di x , l'uguaglianza $L(e^{\lambda x}) = 0$ è soddisfatta se e solo se il polinomio è nullo:

$$L(e^{\lambda x}) = 0 \iff p(\lambda) = 0$$

□

E' necessario studiare l'equazione caratteristica anche in campo complesso per i seguenti motivi:

1. **Esistenza delle soluzioni:** il teorema fondamentale dell'algebra garantisce che ogni equazione polinomiale ha sempre soluzioni nel campo complesso \mathbb{C} , mentre nel campo reale \mathbb{R} alcune equazioni potrebbero non avere soluzioni;
2. **Completezza della soluzione:** per costruire la soluzione generale completa dell'EDO, abbiamo bisogno di tutte le radici dell'equazione caratteristica, incluse quelle complesse.

Considerando l'equazione caratteristica quadratica:

$$\lambda^2 + b\lambda + c = 0$$

Possiamo dire che per il teorema fondamentale dell'algebra ci viene garantito che esistono sempre esattamente 2 radici in \mathbb{C} (contate con molteplicità). Più semplicemente, il teorema garantisce che non possiamo mai avere assenza di soluzioni: l'equazione caratteristica ha sempre radici, anche se potrebbero essere complesse.

Teorema 1.5.2 (Integrale generale della EDO lineare del secondo ordine omogenea associata a coefficienti costanti). *L'integrale generale della EDO lineare del secondo ordine omogenea associata a coefficienti costanti $ay''(x) + by'(x) + cy(x) = f(x)$ è data da*

$$y_{om}(x, c_1, c_2) = c_1 y_1(x) + c_2 y_2(x), \quad c_1, c_2 \in \mathbb{R}$$

$y_1(x), y_2(x)$ sono determinate dai tre casi distinti.

DIMOSTRAZIONE

Caso 1 $\Delta > 0$ (Due radici reali distinte)

In questo caso il discriminante è positivo:

$$\Delta = b^2 - 4ac > 0,$$

quindi le radici dell'equazione caratteristica $a\lambda^2 + b\lambda + c = 0$ sono due valori reali distinti: $\lambda_1 \neq \lambda_2$, dove $\lambda_1, \lambda_2 \in \mathbb{R}$.

Le soluzioni della EDO lineare omogenea di secondo grado a coefficienti costanti

$$ay''(x) + by'(x) + cy(x) = 0$$

sono

$$y_1(x) = e^{\lambda_1 x} \quad \text{e} \quad y_2(x) = e^{\lambda_2 x}.$$

Passo 1: Verifica dell'indipendenza lineare tramite Wronskiana.

Per dimostrare che $y_1(x)$ e $y_2(x)$ formano un sistema fondamentale di soluzioni, dobbiamo verificare che siano linearmente indipendenti. Questo si fa calcolando il determinante della matrice Wronskiana:

$$W(y_1(x), y_2(x)) = \begin{vmatrix} y_1(x) & y_2(x) \\ y'_1(x) & y'_2(x) \end{vmatrix}.$$

Calcoliamo le derivate:

$$y'_1(x) = \lambda_1 e^{\lambda_1 x}, \quad y'_2(x) = \lambda_2 e^{\lambda_2 x}.$$

Sostituiamo nella Wronskiana:

$$W(y_1(x), y_2(x)) = \begin{vmatrix} e^{\lambda_1 x} & e^{\lambda_2 x} \\ \lambda_1 e^{\lambda_1 x} & \lambda_2 e^{\lambda_2 x} \end{vmatrix}.$$

Calcoliamo il determinante usando la regola $\det = ad - bc$:

$$\begin{aligned} W(y_1(x), y_2(x)) &= e^{\lambda_1 x} \cdot \lambda_2 e^{\lambda_2 x} - e^{\lambda_2 x} \cdot \lambda_1 e^{\lambda_1 x} \\ &= \lambda_2 e^{\lambda_1 x + \lambda_2 x} - \lambda_1 e^{\lambda_1 x + \lambda_2 x} \\ &= \lambda_2 e^{(\lambda_1 + \lambda_2)x} - \lambda_1 e^{(\lambda_1 + \lambda_2)x} \\ &= (\lambda_2 - \lambda_1) e^{(\lambda_1 + \lambda_2)x}. \end{aligned}$$

Poiché $\lambda_1 \neq \lambda_2$ (per $\Delta > 0$), abbiamo $\lambda_2 - \lambda_1 \neq 0$, e poiché $e^{(\lambda_1 + \lambda_2)x} > 0$ per ogni x , concludiamo che

$$W(y_1(x), y_2(x)) \neq 0.$$

Quindi $y_1(x)$ e $y_2(x)$ sono linearmente indipendenti.

Passo 2: Dimostrazione che ogni soluzione è combinazione lineare di $y_1(x)$ e $y_2(x)$.

Sia $y_{\text{om}}(x)$ una soluzione qualsiasi dell'equazione omogenea. Scriviamo $y_{\text{om}}(x)$ nella forma

$$y_{\text{om}}(x) = e^{\lambda_1 x} u(x),$$

dove $u(x)$ è una funzione da determinare.

Poiché $y_{\text{om}}(x)$ è soluzione dell'equazione differenziale, deve soddisfare:

$$a(e^{\lambda_1 x} u(x))'' + b(e^{\lambda_1 x} u(x))' + c(e^{\lambda_1 x} u(x)) = 0.$$

Calcoliamo la prima derivata usando la regola del prodotto:

$$(e^{\lambda_1 x} u(x))' = (e^{\lambda_1 x})' u(x) + e^{\lambda_1 x} u'(x) = \lambda_1 e^{\lambda_1 x} u(x) + e^{\lambda_1 x} u'(x).$$

Calcoliamo la seconda derivata, applicando nuovamente la regola del prodotto:

$$\begin{aligned} (e^{\lambda_1 x} u(x))'' &= (\lambda_1 e^{\lambda_1 x} u(x) + e^{\lambda_1 x} u'(x))' \\ &= (\lambda_1 e^{\lambda_1 x} u(x))' + (e^{\lambda_1 x} u'(x))' \\ &= \lambda_1^2 e^{\lambda_1 x} u(x) + \lambda_1 e^{\lambda_1 x} u'(x) + \lambda_1 e^{\lambda_1 x} u'(x) + e^{\lambda_1 x} u''(x) \\ &= \lambda_1^2 e^{\lambda_1 x} u(x) + 2\lambda_1 e^{\lambda_1 x} u'(x) + e^{\lambda_1 x} u''(x). \end{aligned}$$

Sostituiamo nell'equazione differenziale:

$$a[\lambda_1^2 e^{\lambda_1 x} u(x) + 2\lambda_1 e^{\lambda_1 x} u'(x) + e^{\lambda_1 x} u''(x)] + b[\lambda_1 e^{\lambda_1 x} u(x) + e^{\lambda_1 x} u'(x)] + c[e^{\lambda_1 x} u(x)] = 0.$$

Distribuiamo i coefficienti:

$$a\lambda_1^2 e^{\lambda_1 x} u(x) + 2a\lambda_1 e^{\lambda_1 x} u'(x) + ae^{\lambda_1 x} u''(x) + b\lambda_1 e^{\lambda_1 x} u(x) + be^{\lambda_1 x} u'(x) + ce^{\lambda_1 x} u(x) = 0.$$

Mettiamo in evidenza $e^{\lambda_1 x}$ (che è sempre > 0):

$$e^{\lambda_1 x} [(a\lambda_1^2 + b\lambda_1 + c)u(x) + au''(x) + (2a\lambda_1 + b)u'(x)] = 0.$$

Dividendo per $e^{\lambda_1 x}$:

$$(a\lambda_1^2 + b\lambda_1 + c)u(x) + au''(x) + (2a\lambda_1 + b)u'(x) = 0.$$

Ora, λ_1 è soluzione dell'equazione caratteristica $a\lambda^2 + b\lambda + c = 0$, quindi

$$a\lambda_1^2 + b\lambda_1 + c = 0.$$

Pertanto l'equazione si riduce a:

$$au''(x) + (2a\lambda_1 + b)u'(x) = 0.$$

Dividiamo per a (con $a \neq 0$):

$$u''(x) + \left(2\lambda_1 + \frac{b}{a}\right)u'(x) = 0.$$

Dall'equazione caratteristica $a\lambda^2 + b\lambda + c = 0$, dividendo per a :

$$\lambda^2 + \frac{b}{a}\lambda + \frac{c}{a} = 0.$$

Per le formule di Viète (o fattorizzando), sappiamo che se λ_1 e λ_2 sono le due radici:

$$\lambda^2 + \frac{b}{a}\lambda + \frac{c}{a} = (\lambda - \lambda_1)(\lambda - \lambda_2).$$

Espandendo il prodotto a destra:

$$(\lambda - \lambda_1)(\lambda - \lambda_2) = \lambda^2 - \lambda_2\lambda - \lambda_1\lambda + \lambda_1\lambda_2 = \lambda^2 - (\lambda_1 + \lambda_2)\lambda + \lambda_1\lambda_2.$$

Uguagliando i coefficienti:

$$\lambda_1 + \lambda_2 = -\frac{b}{a}.$$

Quindi:

$$2\lambda_1 + \frac{b}{a} = 2\lambda_1 - (\lambda_1 + \lambda_2) = 2\lambda_1 - \lambda_1 - \lambda_2 = \lambda_1 - \lambda_2.$$

L'equazione diventa:

$$u''(x) + (\lambda_1 - \lambda_2)u'(x) = 0.$$

Riscriviamo come:

$$u''(x) - (\lambda_2 - \lambda_1)u'(x) = 0.$$

Poniamo $v(x) = u'(x)$. Allora $v'(x) = u''(x)$, e l'equazione diventa:

$$v'(x) - (\lambda_2 - \lambda_1)v(x) = 0.$$

Questa è una EDO del primo ordine a variabili separabili. Poniamo $k = \lambda_2 - \lambda_1$:

$$v'(x) - kv(x) = 0.$$

Riscriviamo:

$$v'(x) = kv(x).$$

Separando le variabili (per $v(x) \neq 0$):

$$\frac{dv}{v} = k dx.$$

Integriamo entrambi i membri:

$$\int \frac{dv}{v} = \int k dx \implies \ln|v| = kx + C_0$$

$$|v| = e^{kx+C_0} = e^{C_0} \cdot e^{kx}.$$

Ponendo $c = \pm e^{C_0}$:

$$v(x) = c \cdot e^{kx} = c \cdot e^{(\lambda_2 - \lambda_1)x}.$$

Ricordando che $v(x) = u'(x)$:

$$u'(x) = c \cdot e^{(\lambda_2 - \lambda_1)x}.$$

Integriamo per trovare $u(x)$:

$$u(x) = \int c \cdot e^{(\lambda_2 - \lambda_1)x} dx.$$

Calcoliamo l'integrale (ricordando che $\int e^{ax} dx = \frac{1}{a}e^{ax} + C$):

$$u(x) = c \cdot \frac{1}{\lambda_2 - \lambda_1} e^{(\lambda_2 - \lambda_1)x} + c_2.$$

Poniamo $c_1 = \frac{c}{\lambda_2 - \lambda_1}$:

$$u(x) = c_1 \cdot e^{(\lambda_2 - \lambda_1)x} + c_2, \quad c_1, c_2 \in \mathbb{R}.$$

Sostituiamo in $y_{\text{om}}(x) = e^{\lambda_1 x} u(x)$:

$$\begin{aligned} y_{\text{om}}(x) &= e^{\lambda_1 x} [c_1 \cdot e^{(\lambda_2 - \lambda_1)x} + c_2] \\ &= c_1 \cdot e^{\lambda_1 x} \cdot e^{(\lambda_2 - \lambda_1)x} + c_2 \cdot e^{\lambda_1 x} \\ &= c_1 \cdot e^{\lambda_1 x + \lambda_2 x - \lambda_1 x} + c_2 \cdot e^{\lambda_1 x} \\ &= c_1 \cdot e^{\lambda_2 x} + c_2 \cdot e^{\lambda_1 x}. \end{aligned}$$

Quindi ogni soluzione $y_{\text{om}}(x)$ è una combinazione lineare di $y_1(x) = e^{\lambda_1 x}$ e $y_2(x) = e^{\lambda_2 x}$.

Caso 2: $\Delta = 0$ (Radice reale doppia)

In questo caso il discriminante è nullo:

$$\Delta = b^2 - 4ac = 0,$$

quindi l'equazione caratteristica ha una radice doppia:

$$\lambda_1 = \lambda_2 = \lambda = -\frac{b}{2a}.$$

Le soluzioni sono

$$y_1(x) = e^{\lambda x}, \quad y_2(x) = xe^{\lambda x}.$$

Passo 1: Verifica dell'indipendenza lineare tramite Wronskiana.

Calcoliamo le derivate:

$$\begin{aligned} y'_1(x) &= \lambda e^{\lambda x}, \\ y'_2(x) &= (xe^{\lambda x})' = (x)'e^{\lambda x} + x(e^{\lambda x})' = e^{\lambda x} + x\lambda e^{\lambda x} = e^{\lambda x}(1 + \lambda x). \end{aligned}$$

Calcoliamo la Wronskiana:

$$W(y_1(x), y_2(x)) = \begin{vmatrix} e^{\lambda x} & xe^{\lambda x} \\ \lambda e^{\lambda x} & e^{\lambda x}(1 + \lambda x) \end{vmatrix}.$$

Calcoliamo il determinante:

$$\begin{aligned} W(y_1(x), y_2(x)) &= e^{\lambda x} \cdot e^{\lambda x}(1 + \lambda x) - xe^{\lambda x} \cdot \lambda e^{\lambda x} \\ &= e^{2\lambda x}(1 + \lambda x) - \lambda x e^{2\lambda x} \\ &= e^{2\lambda x} + \lambda x e^{2\lambda x} - \lambda x e^{2\lambda x} \\ &= e^{2\lambda x}. \end{aligned}$$

Poiché $e^{2\lambda x} > 0$ per ogni x , concludiamo che

$$W(y_1(x), y_2(x)) \neq 0.$$

Quindi $y_1(x)$ e $y_2(x)$ sono linearmente indipendenti.

Passo 2: Dimostrazione che ogni soluzione è combinazione lineare di y_1 e y_2 .

Sia $y_{\text{om}}(x)$ una soluzione qualsiasi della forma

$$y_{\text{om}}(x) = e^{\lambda x} u(x).$$

Sostituiamo nell'equazione differenziale:

$$a(e^{\lambda x} u(x))'' + b(e^{\lambda x} u(x))' + ce^{\lambda x} u(x) = 0.$$

Calcoliamo la prima derivata:

$$(e^{\lambda x} u(x))' = \lambda e^{\lambda x} u(x) + e^{\lambda x} u'(x).$$

Calcoliamo la seconda derivata:

$$\begin{aligned} (e^{\lambda x} u(x))'' &= (\lambda e^{\lambda x} u(x) + e^{\lambda x} u'(x))' \\ &= \lambda^2 e^{\lambda x} u(x) + \lambda e^{\lambda x} u'(x) + \lambda e^{\lambda x} u'(x) + e^{\lambda x} u''(x) \\ &= \lambda^2 e^{\lambda x} u(x) + 2\lambda e^{\lambda x} u'(x) + e^{\lambda x} u''(x). \end{aligned}$$

Sostituiamo nell'equazione:

$$a[\lambda^2 e^{\lambda x} u(x) + 2\lambda e^{\lambda x} u'(x) + e^{\lambda x} u''(x)] + b[\lambda e^{\lambda x} u(x) + e^{\lambda x} u'(x)] + ce^{\lambda x} u(x) = 0.$$

Mettiamo in evidenza $e^{\lambda x}$:

$$e^{\lambda x}[(a\lambda^2 + b\lambda + c)u(x) + au''(x) + (2a\lambda + b)u'(x)] = 0.$$

Poiché λ è soluzione dell'equazione caratteristica, $a\lambda^2 + b\lambda + c = 0$, quindi:

$$au''(x) + (2a\lambda + b)u'(x) = 0.$$

Dividiamo per a :

$$u''(x) + \left(2\lambda + \frac{b}{a}\right)u'(x) = 0.$$

Poiché $\lambda = -\frac{b}{2a}$:

$$2\lambda = 2 \cdot \left(-\frac{b}{2a}\right) = -\frac{b}{a}.$$

Quindi:

$$2\lambda + \frac{b}{a} = -\frac{b}{a} + \frac{b}{a} = 0.$$

L'equazione diventa:

$$u''(x) = 0.$$

Poniamo $v(x) = u'(x)$. Allora:

$$v'(x) = 0.$$

Integrando:

$$v(x) = c, \quad c \in \mathbb{R}.$$

Quindi $u'(x) = c$. Integriamo nuovamente:

$$u(x) = \int c \, dx = cx + c_2 = c_2 + c_1 x,$$

dove abbiamo rinominato $c = c_1$ e la costante di integrazione è c_2 .

Quindi:

$$u(x) = c_1 + c_2 x, \quad c_1, c_2 \in \mathbb{R}.$$

Sostituiamo in $y_{\text{om}}(x) = e^{\lambda x}u(x)$:

$$y_{\text{om}}(x) = e^{\lambda x}(c_1 + c_2x) = c_1e^{\lambda x} + c_2xe^{\lambda x}.$$

Quindi ogni soluzione $y_{\text{om}}(x)$ è una combinazione lineare di $y_1(x) = e^{\lambda x}$ e $y_2(x) = xe^{\lambda x}$.

Caso 3: $\Delta < 0$ (Due radici complesse coniugate)

In questo caso il discriminante è negativo:

$$\Delta = b^2 - 4ac < 0,$$

quindi l'equazione caratteristica ha due radici complesse coniugate:

$$\lambda_{1,2} = \alpha \pm i\beta,$$

dove

$$\alpha = -\frac{b}{2a}, \quad \beta = \frac{\sqrt{-\Delta}}{2a} = \frac{\sqrt{4ac - b^2}}{2a} \in \mathbb{R}.$$

Le soluzioni complesse sono

$$y_1(x) = e^{\lambda_1 x} = e^{(\alpha+i\beta)x}, \quad y_2(x) = e^{\lambda_2 x} = e^{(\alpha-i\beta)x}.$$

Passo 1: Verifica dell'indipendenza lineare tramite Wronskiana.

Calcoliamo le derivate:

$$y'_1(x) = \lambda_1 e^{\lambda_1 x}, \quad y'_2(x) = \lambda_2 e^{\lambda_2 x}.$$

Calcoliamo la Wronskiana:

$$W(y_1(x), y_2(x)) = \begin{vmatrix} e^{\lambda_1 x} & e^{\lambda_2 x} \\ \lambda_1 e^{\lambda_1 x} & \lambda_2 e^{\lambda_2 x} \end{vmatrix}.$$

Calcoliamo il determinante:

$$\begin{aligned} W(y_1(x), y_2(x)) &= e^{\lambda_1 x} \cdot \lambda_2 e^{\lambda_2 x} - e^{\lambda_2 x} \cdot \lambda_1 e^{\lambda_1 x} \\ &= \lambda_2 e^{(\lambda_1+\lambda_2)x} - \lambda_1 e^{(\lambda_1+\lambda_2)x} \\ &= (\lambda_2 - \lambda_1) e^{(\lambda_1+\lambda_2)x}. \end{aligned}$$

Calcoliamo $\lambda_1 + \lambda_2$ e $\lambda_2 - \lambda_1$:

$$\lambda_1 + \lambda_2 = (\alpha + i\beta) + (\alpha - i\beta) = 2\alpha,$$

$$\lambda_2 - \lambda_1 = (\alpha - i\beta) - (\alpha + i\beta) = -2i\beta.$$

Quindi:

$$W(y_1(x), y_2(x)) = -2i\beta \cdot e^{2\alpha x} \neq 0,$$

poiché $\beta \neq 0$ (essendo $\Delta < 0$).

Quindi $y_1(x)$ e $y_2(x)$ sono linearmente indipendenti.

Passo 2: Dimostrazione che ogni soluzione è combinazione lineare di y_1 e y_2 .

Sia $y_{\text{om}}(x)$ una soluzione qualsiasi della forma

$$y_{\text{om}}(x) = e^{\alpha x}u(x).$$

Sostituiamo nell'equazione differenziale:

$$a[e^{\alpha x}u(x)]'' + b[e^{\alpha x}u(x)]' + ce^{\alpha x}u(x) = 0.$$

Calcoliamo la prima derivata:

$$[e^{\alpha x}u(x)]' = \alpha e^{\alpha x}u(x) + e^{\alpha x}u'(x).$$

Calcoliamo la seconda derivata:

$$\begin{aligned}[e^{\alpha x} u(x)]'' &= [\alpha e^{\alpha x} u(x) + e^{\alpha x} u'(x)]' \\ &= \alpha^2 e^{\alpha x} u(x) + \alpha e^{\alpha x} u'(x) + \alpha e^{\alpha x} u'(x) + e^{\alpha x} u''(x) \\ &= \alpha^2 e^{\alpha x} u(x) + 2\alpha e^{\alpha x} u'(x) + e^{\alpha x} u''(x).\end{aligned}$$

Sostituiamo nell'equazione:

$$a[\alpha^2 e^{\alpha x} u(x) + 2\alpha e^{\alpha x} u'(x) + e^{\alpha x} u''(x)] + b[\alpha e^{\alpha x} u(x) + e^{\alpha x} u'(x)] + c e^{\alpha x} u(x) = 0.$$

Mettiamo in evidenza $e^{\alpha x}$:

$$e^{\alpha x} [au''(x) + (2a\alpha + b)u'(x) + (a\alpha^2 + b\alpha + c)u(x)] = 0.$$

Dividiamo per $e^{\alpha x}$ e per a :

$$u''(x) + \left(2\alpha + \frac{b}{a}\right)u'(x) + \left(\alpha^2 + \frac{b\alpha}{a} + \frac{c}{a}\right)u(x) = 0.$$

Sostituiamo $\alpha = -\frac{b}{2a}$:

$$2\alpha + \frac{b}{a} = 2 \cdot \left(-\frac{b}{2a}\right) + \frac{b}{a} = -\frac{b}{a} + \frac{b}{a} = 0.$$

Calcoliamo il coefficiente di $u(x)$:

$$\begin{aligned}\alpha^2 + \frac{b\alpha}{a} + \frac{c}{a} &= \left(-\frac{b}{2a}\right)^2 + \frac{b}{a} \cdot \left(-\frac{b}{2a}\right) + \frac{c}{a} \\ &= \frac{b^2}{4a^2} - \frac{b^2}{2a^2} + \frac{c}{a} \\ &= \frac{b^2 - 2b^2 + 4ac}{4a^2} = \frac{-b^2 + 4ac}{4a^2}.\end{aligned}$$

Poiché $\Delta = b^2 - 4ac$:

$$-b^2 + 4ac = -\Delta.$$

Quindi:

$$\alpha^2 + \frac{b\alpha}{a} + \frac{c}{a} = \frac{-\Delta}{4a^2}.$$

Ricordando che $\beta = \frac{\sqrt{-\Delta}}{2a}$, abbiamo:

$$\beta^2 = \left(\frac{\sqrt{-\Delta}}{2a}\right)^2 = \frac{-\Delta}{4a^2}.$$

L'equazione diventa:

$$u''(x) + \beta^2 u(x) = 0.$$

Questa è l'equazione dell'oscillatore armonico, la cui soluzione generale è:

$$u(x) = c_1 \cos(\beta x) + c_2 \sin(\beta x), \quad c_1, c_2 \in \mathbb{R}.$$

Sostituiamo in $y_{\text{om}}(x) = e^{\alpha x} u(x)$:

$$y_{\text{om}}(x) = e^{\alpha x} [c_1 \cos(\beta x) + c_2 \sin(\beta x)].$$

Quindi l'integrale generale in forma reale è:

$$y_{\text{om}}(x) = e^{\alpha x} [c_1 \cos(\beta x) + c_2 \sin(\beta x)].$$

Questo può essere scritto come combinazione lineare di:

$$y_1(x) = e^{\alpha x} \cos(\beta x), \quad y_2(x) = e^{\alpha x} \sin(\beta x).$$

□

1.5.3 Struttura di una curva integrale di una EDO lineare del secondo ordine completa a coefficienti costanti

Ci sono tre metodi per la ricerca di una curva integrale di una EDO lineare del secondo ordine completa a coefficienti costanti di partenza in modo da verificare completamente le condizioni iniziali: ad occhio, con il metodo di variazione delle costanti e con il metodo di somiglianza.

Metodo di somiglianza

Si analizza le fattezze del termine noto per intuire le soluzioni particolari tramite una tabella in cui è presente un riepilogo di tutti i casi dove si trova la soluzione per similitudine.

Data una EDO lineare del secondo ordine a coefficienti costanti $ay''(x) + by'(x) + cy(x) = f(x)$ è necessario distinguere i seguenti casi

Termine noto $f(x)$	Soluzione $y_p(x)$	Modifiche necessarie
$e^{\alpha x}$	$Ae^{\alpha x}$	• Moltiplicare per x se α è radice semplice, moltiplicare per x^2 se α è radice doppia
$\cos(\beta x)$ o $\sin(\beta x)$	$A \cos(\beta x) + B \sin(\beta x)$	• Moltiplicare per x se $i\beta$ è radice
$P_n(x)$ (polinomio grado n)	$Q_n(x)$ (polinomio grado n)	• Moltiplicare per x se 0 è radice semplice, moltiplicare per x^2 se 0 è radice doppia
$e^{\alpha x} P_n(x)$	$e^{\alpha x} Q_n(x)$	• Applicare le regole come sopra per α
$e^{\alpha x} \cos(\beta x)$ o $e^{\alpha x} \sin(\beta x)$	$e^{\alpha x} [A \cos(\beta x) + B \sin(\beta x)]$	• Moltiplicare per x se $\alpha \pm i\beta$ è radice

Metodo di variazione delle costanti

Il metodo delle variazioni delle costanti ha la soluzione particolare della EDO lineare del secondo ordine a coefficienti costanti della forma $\bar{y}(x) = c_1(x)y_1(x) + c_2(x)y_2(x)$ dove $x \in I = [a, b]$ e $c_{1,2} \in C^2(I)$ sono da determinare.

La formula che si applica per la soluzione particolare della EDO del primo ordine si estende di una dimensione per l'aggiunta $c_2(x)y_2(x)$.

Inoltre, imponendo che la funzione della forma $\bar{y}(x) = c_1(x)y_1(x) + c_2(x)y_2(x)$ sia soluzione particolare, sarà ottenuto un sistema nelle ipotesi c'_1 e c'_2 che sarà risolto tramite la regola di Cramer (teoria dei sistemi lineari).

DIMOSTRAZIONE

Consideriamo la EDO lineare del secondo ordine non omogenea a coefficienti costanti:

$$ay''(x) + by'(x) + cy(x) = f(x), \quad a \neq 0$$

Vogliamo trovare una soluzione particolare $\bar{y}(x)$ di questa equazione. Quindi, supponiamo di conoscere due soluzioni linearmente indipendenti $y_1(x)$ e $y_2(x)$ dell'equazione omogenea associata e cerchiamo una soluzione particolare della forma:

$$\bar{y}(x) = c_1(x)y_1(x) + c_2(x)y_2(x)$$

dove $c_1(x)$ e $c_2(x)$ sono funzioni incognite da determinare (non più costanti come nell'omogenea).

Passo 1: Deriviamo $\bar{y}(x)$ rispetto a x applicando la regola del prodotto.

$$\bar{y}'(x) = \frac{d}{dx}[c_1(x)y_1(x)] + \frac{d}{dx}[c_2(x)y_2(x)]$$

$$\bar{y}'(x) = c'_1(x)y_1(x) + c_1(x)y'_1(x) + c'_2(x)y_2(x) + c_2(x)y'_2(x)$$

Passo 2: Prima condizione (trucco di Lagrange).

Per semplificare i calcoli successivi, imponiamo la condizione:

$$c'_1(x)y_1(x) + c'_2(x)y_2(x) = 0$$

Questa è una scelta arbitraria che ci permette di eliminare alcuni termini. La derivata prima diventa:

$$\bar{y}'(x) = c_1(x)y'_1(x) + c_2(x)y'_2(x)$$

Passo 3: Calcolo della derivata seconda. Deriviamo nuovamente $\bar{y}'(x)$ applicando la regola del prodotto:

$$\bar{y}''(x) = \frac{d}{dx} [c_1(x)y'_1(x) + c_2(x)y'_2(x)]$$

$$\bar{y}''(x) = \frac{d}{dx}[c_1(x)y'_1(x)] + \frac{d}{dx}[c_2(x)y'_2(x)]$$

$$\bar{y}''(x) = c'_1(x)y'_1(x) + c_1(x)y''_1(x) + c'_2(x)y'_2(x) + c_2(x)y''_2(x)$$

Passo 4: Sostituzione nell'EDO.

Sostituiamo $\bar{y}(x)$, $\bar{y}'(x)$ e $\bar{y}''(x)$ nell'equazione differenziale originale:

$$ay''(x) + by'(x) + cy(x) = f(x)$$

Sostituiamo:

$$\begin{aligned} a &[c'_1(x)y'_1(x) + c_1(x)y''_1(x) + c'_2(x)y'_2(x) + c_2(x)y''_2(x)] \\ &+ b [c_1(x)y'_1(x) + c_2(x)y'_2(x)] \\ &+ c [c_1(x)y_1(x) + c_2(x)y_2(x)] = f(x) \end{aligned}$$

Passo 5: Riordino dei termini.

Distribuiamo i coefficienti a , b e c :

$$\begin{aligned} ac'_1(x)y'_1(x) &+ ac_1(x)y''_1(x) + ac'_2(x)y'_2(x) + ac_2(x)y''_2(x) \\ &+ bc_1(x)y'_1(x) + bc_2(x)y'_2(x) \\ &+ cc_1(x)y_1(x) + cc_2(x)y_2(x) = f(x) \end{aligned}$$

Raccogliamo i termini con $c_1(x)$ e con $c_2(x)$:

$$\begin{aligned} c_1(x) &[ay''_1(x) + by'_1(x) + cy_1(x)] \\ &+ c_2(x) [ay''_2(x) + by'_2(x) + cy_2(x)] \\ &+ a [c'_1(x)y'_1(x) + c'_2(x)y'_2(x)] = f(x) \end{aligned}$$

Passo 6: Semplificazione.

Poiché $y_1(x)$ e $y_2(x)$ sono soluzioni dell'equazione omogenea, abbiamo:

$$ay''_1(x) + by'_1(x) + cy_1(x) = 0$$

$$ay''_2(x) + by'_2(x) + cy_2(x) = 0$$

Quindi i primi due termini si annullano:

$$c_1(x) \cdot 0 + c_2(x) \cdot 0 + a [c'_1(x)y'_1(x) + c'_2(x)y'_2(x)] = f(x)$$

L'equazione si riduce a:

$$a [c'_1(x)y'_1(x) + c'_2(x)y'_2(x)] = f(x)$$

Dividendo entrambi i membri per $a \neq 0$:

$$c'_1(x)y'_1(x) + c'_2(x)y'_2(x) = \frac{f(x)}{a}$$

Passo 7: Sistema di equazioni.

Otteniamo il sistema lineare 2×2 nelle incognite $c'_1(x)$ e $c'_2(x)$:

$$\begin{cases} c'_1(x)y_1(x) + c'_2(x)y_2(x) = 0 \\ c'_1(x)y'_1(x) + c'_2(x)y'_2(x) = \frac{f(x)}{a} \end{cases}$$

Passo 8: Verifica dell'esistenza e unicità della soluzione.

La matrice dei coefficienti del sistema è:

$$A = \begin{pmatrix} y_1(x) & y_2(x) \\ y'_1(x) & y'_2(x) \end{pmatrix}$$

Il determinante di questa matrice è:

$$\det(A) = \begin{vmatrix} y_1(x) & y_2(x) \\ y'_1(x) & y'_2(x) \end{vmatrix} = y_1(x)y'_2(x) - y_2(x)y'_1(x) = W(x) \neq 0$$

Poiché $W(x) \neq 0$ (le soluzioni y_1 e y_2 sono linearmente indipendenti), il sistema ammette un'unica soluzione per le incognite $c'_1(x)$ e $c'_2(x)$.

Passo 9: Risoluzione del sistema con la regola di Cramer.

Applichiamo la regola di Cramer per trovare $c'_1(x)$ e $c'_2(x)$.

Per $c'_1(x)$:

Secondo la regola di Cramer, $c'_1(x)$ si ottiene sostituendo la prima colonna della matrice dei coefficienti con il vettore dei termini noti e dividendo per il determinante $W(x)$:

$$c'_1(x) = \frac{\begin{vmatrix} 0 & y_2(x) \\ \frac{f(x)}{a} & y'_2(x) \end{vmatrix}}{W(x)}$$

Calcoliamo il determinante al numeratore:

$$\begin{vmatrix} 0 & y_2(x) \\ \frac{f(x)}{a} & y'_2(x) \end{vmatrix} = 0 \cdot y'_2(x) - y_2(x) \cdot \frac{f(x)}{a} = -\frac{y_2(x)f(x)}{a}$$

Quindi:

$$c'_1(x) = \frac{-\frac{y_2(x)f(x)}{a}}{W(x)} = -\frac{y_2(x)f(x)}{a \cdot W(x)}$$

Per $c'_2(x)$:

Secondo la regola di Cramer, $c'_2(x)$ si ottiene sostituendo la seconda colonna della matrice dei coefficienti con il vettore dei termini noti e dividendo per il determinante $W(x)$:

$$c'_2(x) = \frac{\begin{vmatrix} y_1(x) & 0 \\ y'_1(x) & \frac{f(x)}{a} \end{vmatrix}}{W(x)}$$

Calcoliamo il determinante al numeratore:

$$\begin{vmatrix} y_1(x) & 0 \\ y'_1(x) & \frac{f(x)}{a} \end{vmatrix} = y_1(x) \cdot \frac{f(x)}{a} - 0 \cdot y'_1(x) = \frac{y_1(x)f(x)}{a}$$

Quindi:

$$c'_2(x) = \frac{\frac{y_1(x)f(x)}{a}}{W(x)} = \frac{y_1(x)f(x)}{a \cdot W(x)}$$

Passo 10: Integrazione.

Per trovare $c_1(x)$ e $c_2(x)$, integriamo $c'_1(x)$ e $c'_2(x)$ rispetto a x :

$$\begin{aligned} c_1(x) &= \int c'_1(x) dx = \int \left(-\frac{y_2(x)f(x)}{a \cdot W(x)} \right) dx = - \int \frac{y_2(x)f(x)}{a \cdot W(x)} dx \\ c_2(x) &= \int c'_2(x) dx = \int \frac{y_1(x)f(x)}{a \cdot W(x)} dx \end{aligned}$$

Non aggiungiamo costanti di integrazione perché stiamo cercando una soluzione particolare della EDO non omogenea, non la soluzione generale. Le costanti arbitrarie saranno presenti nella soluzione dell'omogenea.

Passo 11: Soluzione particolare.

La soluzione particolare dell'EDO non omogenea è:

$$\bar{y}(x) = c_1(x)y_1(x) + c_2(x)y_2(x)$$

Esplicitamente:

$$\bar{y}(x) = \left[- \int \frac{y_2(x)f(x)}{a \cdot W(x)} dx \right] y_1(x) + \left[\int \frac{y_1(x)f(x)}{a \cdot W(x)} dx \right] y_2(x)$$

□

1.5.4 Problema di Cauchy per EDO lineare del secondo ordine a coefficienti costanti

Per una EDO lineare del II ordine della forma $ay''(x) + by'(x) + cy(x) = f(x)$ con a, b, c costanti, $a \neq 0$ e $f(x)$ è continua su un intervallo I dove è presente x_0

$$\begin{cases} ay''(x) + by'(x) + cy(x) &= f(x) \\ y(x_0) &= y_0 \\ y'(x_0) &= y_1 \end{cases}$$

Come si risolve?

1. Trova l'integrale generale dell'equazione omogenea $y_{om}(x) = c_1y_1(x) + c_2y_2(x)$
2. Trova una soluzione particolare $\bar{y}(x)$ usando uno dei metodi descritti
3. Unendo tutto si ottiene $y(x) = y_{om}(x) + \bar{y}(x) = c_1y_1(x) + c_2y_2(x) + \bar{y}(x)$
4. Bisogna trovare le condizioni iniziali per determinare c_1 e c_2 :
 $y(x_0) = c_1y_1(x_0) + c_2y_2(x_0) + \bar{y}(x_0) = y_0$
 $y'(x_0) = c_1y'_1(x_0) + c_2y'_2(x_0) + \bar{y}'(x_0) = y_1$
5. Si risolve il sistema 2x2 con le incognite c_1 e c_2

Cenni di algebra lineare sugli spazi vettoriali

2.1 Il prodotto scalare e la norma

Definizione 2.1.1 (Prodotto Scalare). *E' una funzione definita come*

$$\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$$

Normalmente, è definito

$$\langle \underline{x}, \underline{y} \rangle := x_1 y_1 + \dots + x_n y_n \in \mathbb{R}, \quad \forall \underline{x}, \underline{y} \in \mathbb{R}^n$$

Le sue proprietà sono:

- **Bilinearità**

$$\langle \alpha \underline{x}_1 + \beta \underline{x}_2, \underline{y} \rangle = \alpha \langle \underline{x}_1, \underline{y} \rangle + \beta \langle \underline{x}_2, \underline{y} \rangle, \quad \forall \underline{x}_1, \underline{x}_2, \underline{y} \in \mathbb{R}^n \text{ e } \alpha, \beta \in \mathbb{R}$$

Vale lo stesso per $\langle \alpha \underline{y}_1 + \beta \underline{y}_2, \underline{x} \rangle$

- **Simmetria**

$$\langle \underline{x}, \underline{y} \rangle = \langle \underline{y}, \underline{x} \rangle, \quad \forall \underline{x}, \underline{y} \in \mathbb{R}^n$$

- **Positività**

$$\langle \underline{x}, \underline{x} \rangle > 0 \text{ oppure } \langle \underline{x}, \underline{x} \rangle = 0 \Leftrightarrow \underline{x} = 0, \quad \forall \underline{x} \in \mathbb{R}^n$$

Definizione 2.1.2 (La lunghezza del vettore: Norma). *Sia $\underline{x} \in \mathbb{R}^n$ un vettore, la sua lunghezza o norma è definita come la radice quadrata del prodotto scalare del vettore con se stesso:*

$$\|\underline{x}\| := \sqrt{\langle \underline{x}, \underline{x} \rangle}$$

Essendo in \mathbb{R}^n la norma soddisfa la proprietà di positività del prodotto scalare, quindi $\|\underline{x}\| : \mathbb{R}^n \rightarrow \mathbb{R}_0^+$.

La radice quadrata risulta importante nella rappresentazione della norma perché siamo in \mathbb{R}^n , di conseguenza, il prodotto scalare di un vettore per se stesso è la somma dei quadrati delle componenti del vettore:

$$\langle \underline{x}, \underline{x} \rangle = x_1^2 + x_2^2 + \dots + x_n^2$$

Il teorema di Pitagora ci fornisce la seguente uguaglianza $a^2 + b^2 = c^2$: bisogna estrarre la radice per trovare la lunghezza effettiva (c) rispettando la proprietà di omogeneità e mantenendo la coerenza con la nostra idea di "scala".

- **Positività**

$$\|\underline{x}\| = 0 \Leftrightarrow \underline{x} = 0, \quad \forall \underline{x} \in \mathbb{R}^n$$

- **Omogeneità**

$$\|\lambda \underline{x}\| = |\lambda| \cdot \|\underline{x}\|, \quad \forall \underline{x} \in \mathbb{R}^n \text{ e } \forall \lambda \in \mathbb{R}$$

- **Disuguaglianza triangolare**

$$\|\underline{x} + \underline{y}\| \leq \|\underline{x}\| + \|\underline{y}\|, \quad \forall \underline{x}, \underline{y} \in \mathbb{R}^n$$

DIMOSTRAZIONE

Sviluppo del quadrato della norma

$$\|x + y\|^2 = \langle x + y, x + y \rangle$$

Proprietà di bilinearità e simmetria

$$\begin{aligned} \|x + y\|^2 &= \langle x, x \rangle + 2\langle x, y \rangle + \langle y, y \rangle \\ \|x + y\|^2 &= \|x\|^2 + 2\langle x, y \rangle + \|y\|^2 \end{aligned}$$

Posso riscrivermelo secondo la seguente disuguaglianza:

$$\|x + y\|^2 \leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2$$

Riconoscimento del quadrato di binomio

$$\|x + y\|^2 \leq (\|x\| + \|y\|)^2$$

Estrazione della radice quadrata (essendo termini non negativi)

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

□

2.1.1 Disuguaglianze fondamentali indotte da una norma

Definizione 2.1.3 (Formula di Carnot). Per ogni $\underline{x}, \underline{y} \in \mathbb{R}^n$, vale:

$$\|\underline{x} + \underline{y}\|^2 = \|\underline{x}\|^2 + \|\underline{y}\|^2 + 2\langle \underline{x}, \underline{y} \rangle$$

La formula di Carnot generalizza il teorema di Pitagora agli spazi vettoriali dotati di prodotto scalare. Essa ci dice come calcolare il quadrato della norma della somma di due vettori in termini delle norme dei singoli vettori e del loro prodotto scalare.

Geometricamente, se pensiamo a \underline{x} e \underline{y} come ai lati di un parallelogramma, $\underline{x} + \underline{y}$ rappresenta la diagonale.

DIMOSTRAZIONE

Partiamo dalla definizione di norma indotta dal prodotto scalare:

$$\begin{aligned} \|\underline{x} + \underline{y}\|^2 &= \langle \underline{x} + \underline{y}, \underline{x} + \underline{y} \rangle \\ &= \langle \underline{x}, \underline{x} \rangle + \langle \underline{x}, \underline{y} \rangle + \langle \underline{y}, \underline{x} \rangle + \langle \underline{y}, \underline{y} \rangle \end{aligned}$$

Per la simmetria del prodotto scalare, $\langle \underline{x}, \underline{y} \rangle = \langle \underline{y}, \underline{x} \rangle$, quindi:

$$\|\underline{x} + \underline{y}\|^2 = \|\underline{x}\|^2 + \|\underline{y}\|^2 + 2\langle \underline{x}, \underline{y} \rangle$$

Caso particolare

Quando i vettori \underline{x} e \underline{y} sono ortogonali, cioè $\langle \underline{x}, \underline{y} \rangle = 0$, la formula di Carnot si riduce al teorema di Pitagora:

$$\|\underline{x} + \underline{y}\|^2 = \|\underline{x}\|^2 + \|\underline{y}\|^2$$

Questa equivalenza è fondamentale:

$$\langle \underline{x}, \underline{y} \rangle = 0 \Leftrightarrow \|\underline{x} + \underline{y}\|^2 = \|\underline{x}\|^2 + \|\underline{y}\|^2$$

□

Teorema 2.1.1 (Disuguaglianza di Cauchy-Schwarz). *Per ogni $\underline{x}, \underline{y} \in \mathbb{R}^n$, vale:*

$$|\langle \underline{x}, \underline{y} \rangle| \leq \|\underline{x}\| \cdot \|\underline{y}\|$$

Inoltre, vale l'uguaglianza se e solo se i vettori sono linearmente dipendenti, ossia:

$$|\langle \underline{x}, \underline{y} \rangle| = \|\underline{x}\| \cdot \|\underline{y}\| \Leftrightarrow \underline{x} \text{ e } \underline{y} \text{ sono linearmente dipendenti}$$

Equivalentemente: $\underline{x} = \underline{0}$ oppure $\underline{y} = \underline{0}$ oppure $\exists \lambda \in \mathbb{R} \setminus \{0\}$ tale che $\underline{x} = \lambda \underline{y}$.

DIMOSTRAZIONE

Caso 1: Se $\underline{y} = \underline{0}$, entrambi i membri dell'disuguaglianza sono nulli e la tesi è verificata.

Caso 2: Supponiamo $\underline{y} \neq \underline{0}$. Consideriamo, per ogni $t \in \mathbb{R}$, il vettore:

$$\underline{z}(t) = \underline{x} + t\underline{y}$$

Per la positività della norma, $\|\underline{z}(t)\|^2 \geq 0$ per ogni t . Calcoliamo usando la formula di Carnot:

$$\begin{aligned} \|\underline{z}(t)\|^2 &= \|\underline{x} + t\underline{y}\|^2 \\ &= \|\underline{x}\|^2 + 2t\langle \underline{x}, \underline{y} \rangle + t^2\|\underline{y}\|^2 \end{aligned}$$

Questa è una funzione quadratica in t della forma $at^2 + bt + c$ con:

- $a = \|\underline{y}\|^2 > 0$ (poiché $\underline{y} \neq \underline{0}$)
- $b = 2\langle \underline{x}, \underline{y} \rangle$
- $c = \|\underline{x}\|^2 \geq 0$

Poiché questa parabola è sempre non negativa (con $a > 0$), il discriminante deve soddisfare:

$$\Delta = b^2 - 4ac \leq 0$$

Sostituendo i valori:

$$\begin{aligned} (2\langle \underline{x}, \underline{y} \rangle)^2 - 4\|\underline{y}\|^2\|\underline{x}\|^2 &\leq 0 \\ 4\langle \underline{x}, \underline{y} \rangle^2 &\leq 4\|\underline{x}\|^2\|\underline{y}\|^2 \\ \langle \underline{x}, \underline{y} \rangle^2 &\leq \|\underline{x}\|^2\|\underline{y}\|^2 \end{aligned}$$

Prendendo la radice quadrata di entrambi i membri (entrambi non negativi):

$$|\langle \underline{x}, \underline{y} \rangle| \leq \|\underline{x}\| \cdot \|\underline{y}\|$$

□

Caso di uguaglianza: L'uguaglianza $|\langle \underline{x}, \underline{y} \rangle| = \|\underline{x}\| \cdot \|\underline{y}\|$ vale se e solo se $\Delta = 0$. C'è uguaglianza se $\underline{y} = \underline{0}$, o se $\underline{x} = \underline{0}$, allora entrambi i membri sono zero oppure $\underline{x} = \lambda \underline{y}$ con $\lambda \in \mathbb{R}$, $\lambda \neq 0$:

$$\begin{aligned} |\langle \underline{x}, \underline{y} \rangle| &= |\langle \lambda \underline{y}, \underline{y} \rangle| = |\lambda| \cdot \|\underline{y}\|^2 \\ \|\underline{x}\| \cdot \|\underline{y}\| &= \|\lambda \underline{y}\| \cdot \|\underline{y}\| = |\lambda| \cdot \|\underline{y}\|^2 \end{aligned}$$

che sono uguali.

□

Nota. Geometricamente, la disuguaglianza di Cauchy-Schwarz afferma che il valore del prodotto scalare tra due vettori non può superare il prodotto delle loro norme. In altre parole, il grado di allineamento tra i vettori, misurato dal prodotto scalare, è al massimo pari al prodotto delle loro lunghezze. L'uguaglianza si verifica soltanto quando i due vettori hanno la stessa direzione (parallelî) oppure direzioni opposte (antiparalleli).

Poiché il prodotto scalare può essere scritto come

$$\langle \underline{x}, \underline{y} \rangle = \|\underline{x}\| \|\underline{y}\| \cos \theta,$$

dove θ è l'angolo compreso tra i due vettori, la diseguaglianza di Cauchy–Schwarz si esprime come

$$\|\underline{x}\| \|\underline{y}\| \cos \theta \leq \|\underline{x}\| \|\underline{y}\|.$$

Dividendo entrambi i membri per $\|\underline{x}\| \|\underline{y}\|$ (supposte non nulle), si ottiene

$$\cos \theta \leq 1,$$

che è sempre vera per ogni angolo reale θ .

L'uguaglianza si verifica quando $|\cos \theta| = 1$:

- $\theta = 0$: vettori paralleli ($\lambda > 0$);
- $\theta = \pi$: vettori antiparalleli ($\lambda < 0$).

2.2 Elementi di topologia indotti da una norma per uno spazio metrico

Dal punto di vista algebrico, \mathbb{R}^n è l'insieme di tutte le ennuple ordinate di numeri reali, ossia un grande contentitore di numeri. Quindi, \mathbb{R}^n è lo spazio dove "vivono" i punti fatti da n coordinate.

Definizione 2.2.1 (Spazio metrico). Una coppia (X, d) è un insieme X a cui si associa una distanza (o metrica) d come $d : X \times X \rightarrow \mathbb{R}_0^+$, ossia ad ogni $(x, y) \in X$ si associa un numero reale non negativo $d(x, y)$

X prenderà la forma di \mathbb{R}^n se si tratta di uno spazio metrico dotato della distanza euclidea. In uno spazio \mathbb{R}^n sono considerati $\underline{x} = \{x_1, x_2, \dots, x_n\} \in \mathbb{R}^n$ vettori, con le seguenti proprietà:

- Moltiplicazione per uno scalare

$$c\underline{x} = \{cx_1, cx_2, \dots, cx_n\} \in \mathbb{R}^n, \quad \forall c \in \mathbb{R} \text{ e } \underline{x} \in \mathbb{R}^n$$

- Somma vettoriale

$$\underline{x} + \underline{y} = (x_1 + y_1, \dots, x_n + y_n), \quad \forall \underline{x}, \underline{y} \in \mathbb{R}^n$$

Norma e Distanza indotta dalla Norma

Definizione 2.2.2 (Norma euclidea). Per ogni $\underline{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$:

$$\|\underline{x}\|_2 := \sqrt{\langle \underline{x}, \underline{x} \rangle} = \sqrt{\sum_{i=1}^n x_i^2}, \quad x_i \in \mathbb{R}_0^+$$

In uno spazio metrico, la distanza e la norma sono due concetti fondamentali che ci permettono di misurare rispettivamente la lunghezza dei vettori e la distanza tra i punti.

- La **Norma** $\|\cdot\| : \mathbb{R}^n \rightarrow \mathbb{R}_0^+$ è una funzione che assegna a ogni vettore un numero reale non negativo e rappresenta la sua lunghezza;
- La **Distanza** (o metrica) $d : X \times X \rightarrow \mathbb{R}_0^+$ è una funzione che misura quanto sono lontani due punti nello spazio.

Definizione 2.2.3 (Distanza indotta dalla Norma). *Sia $\|\cdot\|$ una norma su \mathbb{R}^n . La distanza indotta dalla norma è:*

$$d : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}_0^+ \quad \text{tale che} \quad d(\underline{x}, \underline{y}) = \|\underline{x} - \underline{y}\|$$

Essa eredita le proprietà della norma: positività, simmetria e disugualanza triangolare. Quando consideriamo la norma euclidea $\|\cdot\|_2$, la distanza euclidea diventa:

$$d(\underline{x}, \underline{y}) = \|\underline{x} - \underline{y}\|_2 = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

DIMOSTRAZIONE

Verifichiamo come le proprietà della distanza derivano dalla norma euclidea:

1. **Positività (1)**: essendo $\|\underline{x} - \underline{y}\|_2$ una norma, per definizione è sempre non negativa:

$$d(\underline{x}, \underline{y}) = \|\underline{x} - \underline{y}\|_2 > 0$$

2. **Positività (2)**: la norma è zero se e solo se il vettore è nullo:

$$\begin{aligned} d(\underline{x}, \underline{y}) = 0 &\Leftrightarrow \|\underline{x} - \underline{y}\|_2 = 0 \\ &\Leftrightarrow \underline{x} - \underline{y} = \underline{0} \\ &\Leftrightarrow \underline{x} = \underline{y} \end{aligned}$$

3. **Simmetria**: dalla proprietà di omogeneità della norma $\|\lambda \underline{v}\| = |\lambda| \cdot \|\underline{v}\|$:

$$\begin{aligned} d(\underline{x}, \underline{y}) &= \|\underline{x} - \underline{y}\|_2 \\ &= \|(-1)(\underline{y} - \underline{x})\|_2 \\ &= |-1| \cdot \|\underline{y} - \underline{x}\|_2 \\ &= \|\underline{y} - \underline{x}\|_2 \\ &= d(\underline{y}, \underline{x}) \end{aligned}$$

4. **Disuguaglianza triangolare**: dalla disuguaglianza triangolare della norma $\|\underline{u} + \underline{v}\| \leq \|\underline{u}\| + \|\underline{v}\|$:

$$\begin{aligned} d(\underline{x}, \underline{z}) &= \|\underline{x} - \underline{z}\|_2 \\ &= \|(\underline{x} - \underline{y}) + (\underline{y} - \underline{z})\|_2 \\ &\leq \|\underline{x} - \underline{y}\|_2 + \|\underline{y} - \underline{z}\|_2 \\ &= d(\underline{x}, \underline{y}) + d(\underline{y}, \underline{z}) \end{aligned}$$

□

Definizione 2.2.4 (Distanza Euclidea). *E' una distanza fra due punti che prende la forma*

$$d(\underline{x}, \underline{y}) = \|\underline{x} - \underline{y}\| = \sqrt{\langle \underline{x} - \underline{y}, \underline{y} - \underline{x} \rangle} = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} = \sqrt{(x_1 - y_1)^2 + \dots + (x_n - y_n)^2}$$

$\forall \underline{x}, \underline{y} \in \mathbb{R}^n$

Se $n = 1$, si parla di distanza del tassista (o distanza d_1), che assume la forma

$$d(\underline{x}, \underline{y}) = |x_1 - y_1| + |x_2 - y_2|,$$

dove $\underline{x}, \underline{y} \in \mathbb{R}^2$.

2.2.1 Successione nello spazio metrico

In Analisi 1, una successione in \mathbb{R} è stata definita come una coppia composta da un sottoinsieme di \mathbb{R} e un'applicazione di $\mathbb{N} \rightarrow \mathbb{R}$.

Definizione 2.2.5 (Successione in \mathbb{R}^n). Una successione $\{\underline{x}_n\}$ in \mathbb{R}^n è un sottoinsieme di \mathbb{R}^n del tipo $\underline{x}_n = (\underline{x}_n^1, \dots, \underline{x}_n^n)$, composto da n vettori di n elementi:

$$\underline{x}_n : \mathbb{N} \rightarrow \mathbb{R}^n \quad \text{tale che} \quad n \mapsto \underline{x}_n$$

Quindi, una successione in \mathbb{R}^n è una funzione che associa ad ogni numero naturale n un vettore \underline{x}_n in \mathbb{R}^n . Ogni vettore \underline{x}_n è composto da n componenti, che sono a loro volta numeri reali. In questo modo, possiamo pensare a una successione in \mathbb{R}^n come a una "collezione ordinata" di vettori che evolve al variare di n .

Convergenza attraverso la distanza euclidea

Una successione $\{\underline{x}_n\} \subset \mathbb{R}^n$ con $\underline{x}_n = (\underline{x}_n^1, \dots, \underline{x}_n^n)$ converge ad un vettore $\underline{x} = (x^1, \dots, x^n) \in \mathbb{R}^n$ se la distanza $d(\underline{x}_n, \underline{x}) = \|\underline{x}_n - \underline{x}\| \rightarrow 0$ quando $n \rightarrow +\infty$.

Definizione 2.2.6 (Convergenza). Per quanto sia piccola una distanza ϵ , da un certo punto in poi tutti i termini della successione distano da \underline{x} meno di ϵ .

$$\forall \epsilon > 0 \exists n_\epsilon \in \mathbb{N} \text{ tale che } d(\underline{x}_n, \underline{x}) = \|\underline{x}_n - \underline{x}\| < \epsilon, \forall n \geq n_\epsilon$$

Parlando sempre di convergenza, si può parlare anche di limite convergente:

Definizione 2.2.7 (Limite convergente). Se una successione in \mathbb{R}^n ha un limite che converge \underline{x} , allora le coordinate (o gli elementi) della successione convergono agli elementi di \underline{x} , ossia $\{\underline{x}_n^i\}$ converge a \underline{x}^i . Un limite convergente di una successione in \mathbb{R}^n prende la forma di

$$\lim_{n \rightarrow +\infty} d(\underline{x}_n, \underline{x}) = 0$$

$$\text{con } d(\underline{x}_n, \underline{x}) = \|\underline{x}_n - \underline{x}\| = \sqrt{\sum_{i=1}^n (\underline{x}_n^i - \underline{x}^i)^2} = \sqrt{(\underline{x}_n^1 - \underline{x}^1)^2 + \dots + (\underline{x}_n^n - \underline{x}^n)^2}$$

Quando diciamo che il limite tende a zero quando $n \rightarrow +\infty$ intendiamo formalmente che

$$\forall i = 1..n \quad \underline{x}_n^i \rightarrow \underline{x}^i$$

Tutto ciò significa che $d(\underline{x}_k, \underline{x}) \rightarrow 0$ dove $\underline{x}_k \subset \mathbb{R}^n$ e $\underline{x} \in \mathbb{R}^n$.

Quindi, possiamo affermare che $\{x_k\}$ converge ad x se e solo se ogni componente della successione converge alla corrispondente componente del limite, ossia:

$$\{x_k\} \xrightarrow[k \rightarrow +\infty]{} x \in \mathbb{R}^n \iff \{x_k^i\} \xrightarrow[k \rightarrow +\infty]{} x^i, \quad \forall i = 1, \dots, n$$

La convergenza possiede delle proprietà fondamentali:

- **Unicità**

La successione $\{\underline{x}_n\} \subset \mathbb{R}^n$ ha al più un limite, ossia, se esiste, il limite è unico;

- **Limitatezza**

Se la successione $\{\underline{x}_n\} \subset \mathbb{R}^n$ converge, allora è limitata.

Se i punti si avvicinano sempre più a un limite convergente non possono "scappare all'infinito" ma saranno confinati tutti in una regione limitata dello spazio. Più formalmente:

$$\lim_{n \rightarrow +\infty} d(\underline{x}_n, \underline{x}) = 0 \implies \exists \underline{x}_0 \in \mathbb{R}^n \text{ tale che } d(\underline{x}_n, \underline{x}_0) < M \text{ con } M \in \mathbb{R}$$

- **Sottosuccessione**

Una sottosuccessione $\{\underline{x}\}_{nk} \in \{\underline{x}\}_n$ con $k \in \mathbb{R}$ è convergente a \underline{x} se solo se $\{\underline{x}\}_n$ converge a \underline{x} .

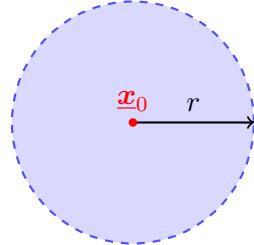
2.2.2 Intorno sferico

Definizione 2.2.8 (Ball/Palla/Disco/Intorno sferico). Sia $\underline{x}_0 \in \mathbb{R}^n$ fissato e $r > 0$ un numero reale positivo, si dice intorno sferico di centro \underline{x}_0 con raggio r l'insieme

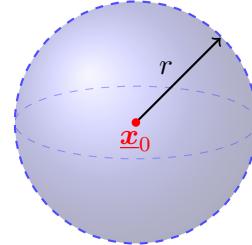
$$B(\underline{x}_0, r) = \{\underline{x} \in \mathbb{R}^n : d(\underline{x}_0, \underline{x}) = \|\underline{x}_0 - \underline{x}\| < r\}$$

Informalmente, un intorno sferico (o "palla") è l'insieme di tutti i punti che stanno dentro una sfera. Immagina di prendere un punto \underline{x}_0 nello spazio, questo è il centro. Poi prendi un numero r , questo è il raggio. La "palla" $B(\underline{x}_0, r)$ contiene tutti i punti che distano meno di r dal centro.

Caso 2D (Cerchio)



Caso 3D (Sfera)



Rappresentazione di un intorno sferico in \mathbb{R}^2 e \mathbb{R}^3

Sia uno spazio metrico $(\mathbb{R}, |\cdot|)$ e siano $\underline{x}_0 \in \mathbb{R}$ e $r > 0$ e $\forall n = 1$

$$\begin{aligned} B(x_0, r) &= \{x \in \mathbb{R} : d(x_0, x) < r\} \\ &= \{x \in \mathbb{R} : |x - x_0| < r\} \\ &= \{x \in \mathbb{R} : x_0 - r < x < x_0 + r\} \\ &= (x_0 - r, x_0 + r) \end{aligned}$$

Sia uno spazio metrico (\mathbb{R}^2, d_2) e sia $\underline{x} = (x_1, x_2)$, siano $\underline{x}_0 \in \mathbb{R}^2$ e $r > 0$

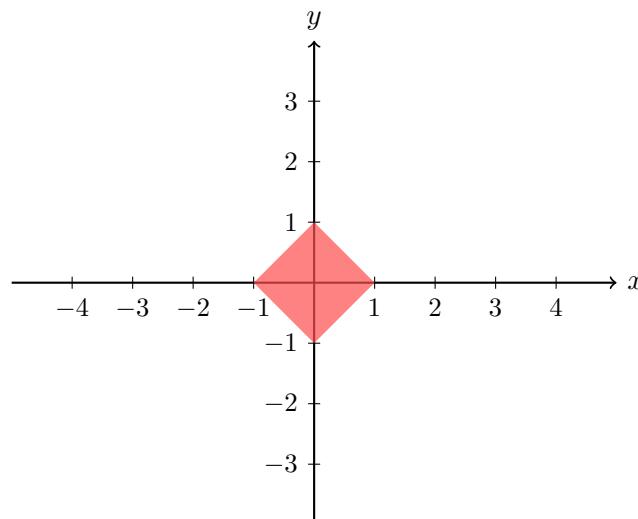
$$\begin{aligned} B(\underline{x}_0, r) &= \{\underline{x} \in \mathbb{R}^2 : d(\underline{x}_0, \underline{x}) < r\} \\ &= \left\{ (x_1, x_2) \in \mathbb{R}^2 : \sqrt{(x_1 - x_0^1)^2 + (x_2 - x_0^2)^2} < r \right\} \end{aligned}$$

Casi di intorno sferico con centro 0 e raggio 1

- Caso della distanza del tassista

Sia uno spazio metrico (\mathbb{R}^2, d_1) e siano $\underline{x}_0 = \underline{0}$ e $r = 1$

$$B(\underline{0}, 1) = \{\underline{x} \in \mathbb{R}^2 : d(\underline{0}, \underline{x}) < 1\} = \{\underline{x} \in \mathbb{R}^2 : |x_1| + |x_2| < 1\}$$

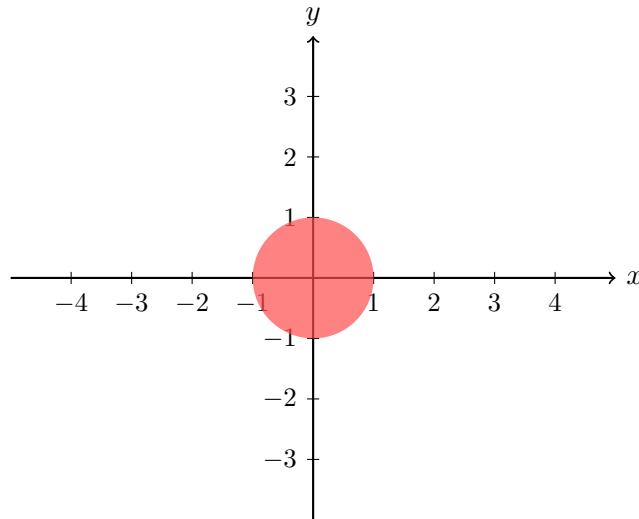


Viene detta distanza del tassista (o distanza di Manhattan), ovvero con $\|\underline{x}\|_1 = 1$, in cui si sommano i valori assoluti di tutte coordinate.

- **Caso della distanza euclidea**

Sia uno spazio metrico (\mathbb{R}^2, d) e siano $\underline{x}_0 = \underline{0}$ e $r = 1$

$$B(\underline{0}, 1) = \{\underline{x} \in \mathbb{R}^2 : d(\underline{0}, \underline{x}) < 1\} = \{\underline{x} \in \mathbb{R}^2 : \sqrt{(x_1 - 0)^2 + (x_2 - 0)^2} < 1\}$$



La forma dell'intorno sferico dipende dalla metrica utilizzata. Con la distanza del tassista, l'intorno sferico è un rombo, mentre con la distanza euclidea è un cerchio.

Nota. Grazie al concetto di intorno sferico si definisce meglio la proprietà della limitatezza. Sia $A \subset \mathbb{R}^n$ e sia lo spazio metrico (\mathbb{R}^n, d) , A si dice limitato se esiste un intorno sferico aperto in cui A è contenuto:

$$\exists r > 0 \text{ e } \underline{x}_0 \in \mathbb{R}^n \text{ tale che } A \subset B(\underline{x}_0, r)$$

Definizione alternativa di successione convergente con intorno sferico

Definizione 2.2.9 (Successione convergente con intorno sferico). Sia una successione $\{x_n\} \subset \mathbb{R}^n$ con $\underline{x}_n = (\underline{x}_n^1, \dots, \underline{x}_n^n)$ che converge ad un vettore $\underline{x} = (\underline{x}^1, \dots, \underline{x}^n) \in \mathbb{R}^n$,

$$\forall \epsilon > 0 \ \exists n_\epsilon \in \mathbb{N} \text{ tale che } \underline{x}_n \in B(\underline{x}, \epsilon), \ \forall n \geq n_\epsilon.$$

2.3 Classificazione topologica degli insiemi

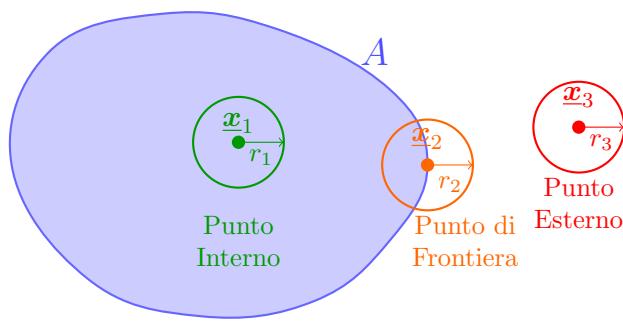
2.3.1 I punti e gli insiemi

Definizione 2.3.1 (Punto Interno). Sia $\underline{x}_0 \in X$ interno ad $A \subset X$ se $\underline{x}_0 \in A$ e esiste almeno un intorno sferico di \underline{x}_0 contenuto in A, ovvero $\exists r > 0$ tale che $B(\underline{x}_0, r) \subset A$.

L'insieme di punti interni di A si scrive $\overset{\circ}{A}$.

Definizione 2.3.2 (Punto Esterno). Sia $\underline{x}_0 \notin X$ esterno ad $A \subset X$ se esiste almeno un intorno sferico disgiunto da A, ovvero $\exists r > 0$ tale che $B(\underline{x}_0, r) \cap A = \emptyset$.

Definizione 2.3.3 (Punto di Frontiera). \underline{x}_0 è un punto di frontiera di A ($\underline{x}_0 \in \delta A$) se ogni intorno di \underline{x}_0 contiene sia punti interni sia punti esterni di A, ossia $\underline{x}_0 \in A$ se $\exists r > 0 : B(\underline{x}_0, r) \cap A \cap A^c \neq \emptyset$ con A^c l'insieme complementare di A tale che $\{\underline{x} \in \mathbb{R}^n : \underline{x} \notin A\}$.



Definizione 2.3.4 (Insieme Aperto). Sia $A \subset X$ se $A = \emptyset$ o se ogni suo punto è interno ad A , ossia $A = \overset{\circ}{A}$.

Formalmente, sia (X, d) uno spazio metrico e sia un sottoinsieme $A \subset X$ si dice aperto se per ogni $\underline{x} \in A$ esiste un raggio $r > 0$ tale che

$$B(\underline{x}, r) \subseteq A.$$

Definizione 2.3.5 (Insieme Chiuso). Un insieme si dice chiuso se contiene tutti i suoi punti di frontiera, ovvero $\delta C \subset C$.

Formalmente, sia (X, d) uno spazio metrico sia un sottoinsieme $C \subseteq X$ si dice chiuso se il suo complementare $X \setminus C$ è un insieme aperto, cioè se per ogni $\underline{x} \in X \setminus C$ esiste un raggio $r > 0$ tale che

$$B(\underline{x}, r) \subseteq X \setminus C.$$

\emptyset e \mathbb{R}^n sono gli unici insiemi aperti e chiusi.

Punto di accumulazione

Definizione 2.3.6 (Punto di accumulazione). Sia $A \subseteq \mathbb{R}^n$ e sia $\underline{x}_0 \in \mathbb{R}^n$ punto di accumulazione per A se $\forall r > 0$ abbiamo $B(\underline{x}_0, r) \cap (A - \{\underline{x}_0\}) \neq \emptyset$.

Preso un intorno sferico, anche piccolissimo, contiene almeno un punto \underline{x} di A diverso dal punto di accumulazione \underline{x}_0

- Sia A aperto, tutti i punti in $\overset{\circ}{A}$ sono punti di accumulazione per A
- Punti di frontiera δA possono essere punti di accumulazione per A

Ogni punto di frontiera che non è punto di accumulazione viene detto punto isolato.

Definizione 2.3.7 (Caratterizzazione sequenziale dei punti di accumulazione). Sia (X, d) uno spazio metrico e sia $A \subseteq X$. Un punto $\underline{x}_0 \in X$ è punto di accumulazione per A se e solo se esiste una successione $\underline{x}_n \subseteq A \setminus \{\underline{x}_0\}$ tale che

$$\lim_{n \rightarrow \infty} \underline{x}_n = \underline{x}_0 \text{ con } n \in \mathbb{N}$$

Un punto è di accumulazione per un insieme se e solo se l'insieme contiene una successione di punti distinti che si avvicinano indefinitamente a esso.

Chiusura di un insieme e dominio

Sia $A \subseteq \mathbb{R}^n$. La chiusura di A è indicata \overline{A} ed è data da:

$$\overline{A} = A \cup A'$$

dove A' è l'insieme dei punti di accumulazione di A .

Quindi, possiamo dire che:

- \bar{A} è chiuso, perché contiene tutti i suoi punti di frontiera;
- È il più piccolo insieme di chiuso di A ;
- Vale sempre $A \subseteq \bar{A}$ e se $A = \bar{A}$ allora A è chiuso.

Definizione 2.3.8 (Dominio in \mathbb{R}^n). Un dominio in \mathbb{R}^n è un insieme che coincide con la chiusura di un insieme aperto A , ossia $A = \bar{A}$.

In questo caso il dominio è un insieme chiuso a cui si aggiunge la frontiera:

$$D = \bar{A} = A \cup \delta A$$

Nota. Un esempio di dominio è una sfera chiusa in \mathbb{R}^n .

Infatti presa un intorno sferico $B(\underline{x}_0, r) = \{\underline{x} \in \mathbb{R}^n : \|\underline{x} - \underline{x}_0\| \leq r\}$ abbiamo che l'insieme di punti interni $\mathring{A} = \{\underline{x} \in \mathbb{R}^n : \|\underline{x} - \underline{x}_0\| < r\}$ è aperto e la frontiera $\delta A = \{\underline{x} \in \mathbb{R}^n : \|\underline{x} - \underline{x}_0\| = r\}$ è la sfera, di conseguenza la sfera è l'unione di A con la sua frontiera ($\bar{B} = \mathring{A} \cup \delta A$).

Una sfera e un punto isolato non formano un dominio perché si contraddice la definizione di dominio che richiede che l'insieme A aperto sia anche connesso, ,in parole povere, non si può dividere A in due insiemi disgiunti.

Funzioni di più variabili

3.1 Definizioni e grafici

Una funzione a due variabili è una legge che associa ad ogni coppia ordinata di numeri reali (x, y) appartenente ad un sottoinsieme $D \subseteq \mathbb{R}^2$ uno ed un solo numero reale $f(x, y)$ che appartiene al codominio.

$$f : D \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$$

Ovvero

$$(x, y) \rightarrow f(x, y)$$

Dove D è l'insieme di definizione, \mathbb{R} è il codominio, (x, y) sono le due variabili indipendenti mentre $f(x, y) = z$ è la variabile dipendente.

Definizione 3.1.1 (Dominio). Il dominio o campo di esistenza è l'insieme di tutti i punti dello spazio per la quale la funzione è matematicamente sensata e restituisce un valore reale.

Trovare il campo di esistenza di una funzione significa trovare l'insieme massimale rispetto all'inclusione su \mathbb{R}^2 .

Immagine di una funzione

L'immagine di una funzione $f : D \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$ è l'insieme di tutti i valori che la funzione assume, o formalmente

$$\text{Im}(f) = \{z \in \mathbb{R} : \exists (x, y) \in D \text{ tale che } z = f(x, y)\}$$

3.1.1 Grafico di funzione e sezioni (trasversali e orizzontali)

Il grafico di una funzione $f : D \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$ è

$$\text{graff} = \{(x, y, z) \in \mathbb{R}^3 : (x, y) \in D, z = f(x, y) \in \mathbb{R}\}$$

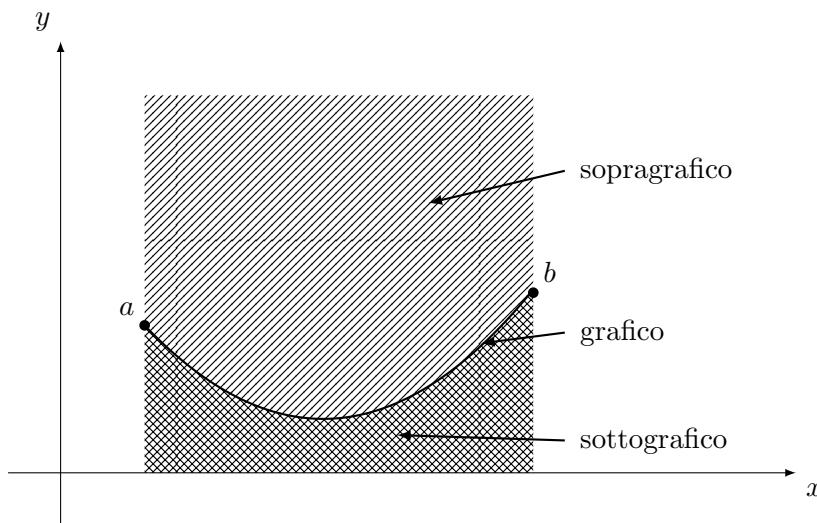
Definizione 3.1.2 (Grafico di una funzione a più variabili). E' un sottoinsieme dello spazio \mathbb{R}^3 , generalmente una superficie nello spazio.

Geometricamente, $\forall (x, y) \in D$ consideriamo un punto $(x, y, f(x, y))$ nello spazio, dove $f(x, y)$ è il valore della funzione.

Nota. Sia $f : I \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$ continua su I e sia il suo grafico $\{(x, y) \in \mathbb{R}^2 : y = f(x)\} \subseteq \mathbb{R}^2$.

Il grafico di f descrive non altro che una curva continua sul piano, definita come un arco di curva continua, e in base alla relazione tra y e $f(x)$ potremmo analizzare una sezione del grafico.

$$\begin{cases} y \leq f(x) & \text{si definisce il sottografico di } f \\ y \geq f(x) & \text{si definisce il sopragrafico di } f \end{cases}$$



Rappresentazione del grafico $y = f(x) \in [a, b]$, diviso in sottografico e sopragrafico.

Per visualizzare o rappresentare il grafico di una funzione $f : D \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$ "seziono" il grafico formando le tracce che intersecano il grafico $\text{graff} \subset \mathbb{R}^3$ con piani nello spazio tridimensionale. Affettando in \mathbb{R}^3 il grafico diminuisco la dimensione del mio problema da \mathbb{R}^3 a \mathbb{R}^2 tracciando delle sezioni del grafico.

Definizione 3.1.3 (Sezione di un grafico). Considerato il grafico di f definita in \mathbb{R}^2 con le terne (x, y, z) , le sezioni piani con $x = k$ o $y = k$ o $z = k$ dove $k \in \mathbb{R}$.

Le due tipologie più importanti di sezioni sono:

- **Sezioni trasversali** ottenute con piani paralleli ai piani coordinati (xz o yz);
- **Sezioni orizzontali** ottentute con piani paralleli al piano xy (con piano di equazione $z = k$).

Le intersezioni con il grafico generalmente producono una curva nello spazio.

Definizione 3.1.4 (Insiemi o curve di livello). Gli insiemi di livello, o anche curve di livello, di una funzione $f : D \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$ sono sottoinsiemi del dominio D su cui la funzione $z = f(x, y)$ assume un valore costante $z = k$.

Formalmente,

$$C_k = \{(x, y) \in D : f(x, y) = k\}$$

Le curve di livello non si intersecano mai (eccetto in punti singolari), perché la funzione non può assumere due valori diversi nello stesso punto.

3.1.2 Curve

Definizione 3.1.5 (Curva). Una curva in \mathbb{R}^n è un'applicazione continua definita su un intervallo $I = [a, b] \subseteq \mathbb{R}$

$$\gamma : I \subseteq \mathbb{R} \rightarrow \mathbb{R}^n$$

Essa è composta da γ_i componenti definite su $I \subseteq \mathbb{R}$ e continue con $i = 1 \dots n$.

Il sostegno o immagine di una curva è l'insieme dei punti geometrici toccati dalla curva nello spazio \mathbb{R}^n

$$\text{Sostegno} = \{x \in \mathbb{R}^n : \exists t \in I, x = \gamma(t)\}$$

Definizione 3.1.6 (Curva Regolare). Sia una curva $\gamma : I \subseteq \mathbb{R} \rightarrow \mathbb{R}^n$ si dice regolare se $\gamma \in C^1(I)$, ossia le sue componenti sono derivabili e continue su I e la derivata $\gamma'(t) \neq 0$.

Una curva regolare non presenta spigoli o cuspidi e in ogni punto è ben definita la retta tangente; infatti con $\gamma'(t) = 0$ la particella si ferma e riparte probabilmente da un'altra direzione creando spigoli.

Definizione 3.1.7 (Curva Cartesiana). Una curva cartesiana è una curva che può essere rappresentata come grafico di una funzione $f : I \subseteq \mathbb{R} \rightarrow \mathbb{R}$, ossia $\gamma(t) = (t, f(t))$.

E' il modo di trasformare il grafico di una funzione ordinaria f in una curva parametrica. Percorriamo il grafico come un cammino usando lo stesso x come parametro.

La prima componente scorre sull'asse x e la seconda componente segue il valore della funzione. La curva cartesiana è un caso particolare di curva regolare, infatti se f è derivabile e $f'(t) \neq 0$ allora $\gamma'(t) = (1, f'(t)) \neq 0$.

Definizione 3.1.8 (Curva Chiusa). Sia una curva $\gamma : I \subseteq \mathbb{R} \rightarrow \mathbb{R}^n$, si dice chiusa se il punto iniziale coincide con quello finale $\gamma(a) = \gamma(b)$.

3.1.3 Accenno al teorema degli zeri per funzioni a più variabili

Teorema 3.1.1 (Teorema degli zeri per funzioni a più variabili). Sia $f : A \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$ continua su A e supponiamo che esistano due punti $P_1 = (x_1, y_1)$ e $P_2 = (x_2, y_2)$ che appartengono ad un insieme connesso $A_i \subseteq A$ tali che $f(P_1) \cdot f(P_2) < 0$, cioè $f(P_1)$ e $f(P_2)$ hanno segno opposto. Allora esiste almeno un punto $P_0 = (x_0, y_0) \in A_i$ tale che $f(P_0) = 0$.

Informalmente, se il grafico di f passa da valori positivi a negativi (o viceversa), deve per forza "attraversare lo zero", ossia intersecare il piano $z = 0$.

3.2 Limiti e continuità

3.2.1 I limiti in più variabili e le loro proprietà

$$f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R} \text{ tale che } \underline{x} \rightarrow f(\underline{x})$$

Definizione 3.2.1 (Limite in \mathbb{R}^n). Sia $\underline{x}_0 = (x_0^1, \dots, x_0^n) \in \mathbb{R}^n$ punto di accumulazione per A si dice che $f(\underline{x}) \rightarrow l$ se $\underline{x} \rightarrow \underline{x}_0$:

$$\lim_{\underline{x} \rightarrow \underline{x}_0} f(\underline{x}) = l$$

Premettiamo che l appartiene all'insieme di numeri reali esteso (\mathbb{R}^*) ad intorni di infinito, semirette del tipo $(-\infty, n)$ e $(n, +\infty)$.

Quindi, $\forall U$ intorno di l in \mathbb{R}^* , esiste un intorno sferico di \underline{x}_0 :

$$\exists r > 0 \text{ tale che } B(\underline{x}_0, r) \subset \mathbb{R}^n \text{ e } f(\underline{x}) \in U \quad \forall \underline{x} \in B(\underline{x}_0, r) \cap (A \setminus \{\underline{x}_0\})$$

Per ogni "vicinanza" del valore limite l , posso scegliere una palla centrata in \underline{x}_0 con raggio abbastanza piccolo in modo che, prendendo qualunque punto \underline{x} del dominio A dentro quella palla (tranne eventualmente il punto \underline{x}_0), il valore $f(\underline{x})$ risulti dentro quella vicinanza di l . In altre parole, se mi avvicino abbastanza a \underline{x}_0 , allora i valori della funzione si avvicinano quanto voglio a l .

Definizione 3.2.2 ($\epsilon - \delta$ con l finito). Se

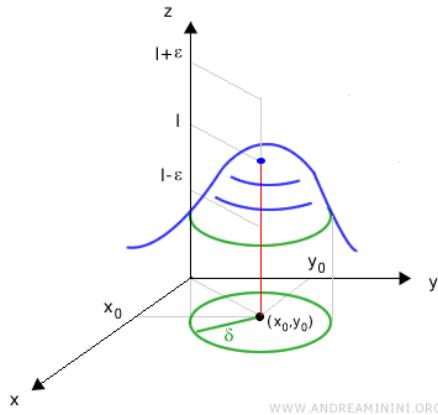
$$\lim_{\underline{x} \rightarrow \underline{x}_0} f(\underline{x}) = l$$

allora

$$\forall \epsilon > 0, \exists \delta > 0 \text{ tale che } |f(\underline{x}) - l| < \epsilon \quad \forall \underline{x} \in (A - \{\underline{x}_0\}) \text{ con } \|\underline{x} - \underline{x}_0\| < \delta$$

che equivale a dire

$$|f(\underline{x}) - l| < \epsilon \iff l - \epsilon < f(\underline{x}) < l + \epsilon \iff f(\underline{x}) \in B(l, \epsilon)$$



WWW.ANDREAMININI.ORG

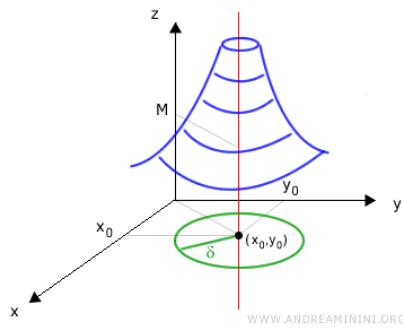
Posso rendere $f(\underline{x})$ vicino a l quando voglio semplicemente prendendo \underline{x} abbastanza vicino a \underline{x}_0 . In altre parole, più ti avvicini a \underline{x}_0 più il valore della funzione si avvicina a l .

Definizione 3.2.3 ($\epsilon - \delta$ con $l = \pm\infty$). Se

$$\lim_{\underline{x} \rightarrow \underline{x}_0} f(\underline{x}) = \pm\infty$$

il limite diverge, allora

$$\forall M > 0, \exists \delta > 0 \text{ tale che } f(\underline{x}) > M \quad \forall \underline{x} \in (A - \{\underline{x}_0\}) \text{ con } \|\underline{x} - \underline{x}_0\| < \delta$$



Posso rendere $f(\underline{x})$ grande quanto voglio semplicemente prendendo \underline{x} abbastanza vicino a \underline{x}_0 . In poche parole, più ti avvicini a \underline{x}_0 più il valore della funzione va verso ∞ . Le proprietà fondamentali del limite:

- **Unicità**

Se esiste, è unico.

Informalmente, se due limiti della stessa funzione esistono finiti e danno lo stesso risultato vuol dire che il limite è necessariamente unico.

Se

$$\lim_{\underline{x}_1 \rightarrow \underline{x}_0} f(\underline{x}_1) = l_1 \text{ e } \lim_{\underline{x}_2 \rightarrow \underline{x}_0} f(\underline{x}_2) = l_2, \quad \text{con } l_1 = l_2$$

allora il limite esiste ed è unico.

- **Composizione**

Il limite di somma o prodotti di funzioni sono uguali alla somma e prodotti dei limiti; vale anche per il quoziente, il limite del rapporto è il rapporto dei limiti, con la condizione che siano ben definiti i limiti (con il denominatore diverso da 0)

Esistenza di un limite

Per esistere un limite deve essere indipendente dalla direzione di avvicinamento di (x, y) al punto di accumulazione \Rightarrow un limite non deve dipendere dal cammino o dalla direzione scelti con il quale (x, y) si avvicina il punto di accumulazione.

Se, scelti due cammini per raggiungere lo stesso punto di accumulazione, il valore dei limiti risulta diverso allora il limite per $f(x, y)$ non esiste.

Definizione 3.2.4 (Limite ristretto ad un sottoinsieme). Sia $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ e sia (x_0, y_0) punto di accumulazione per A . Se

$$\lim_{(x,y) \rightarrow (x_0,y_0)} f(x, y) = l$$

allora per ogni sottoinsieme $C \subseteq A$ tale che (x_0, y_0) sia punto di accumulazione anche per C , vale

$$\lim_{\substack{(x,y) \rightarrow (x_0,y_0) \\ (x,y) \in C}} f(x, y) = l$$

Il limite di f per $(x, y) \rightarrow (x_0, y_0)$ è uguale al limite di f per $(x, y) \rightarrow (x_0, y_0)$ con (x, y) appartenente a qualsiasi sottoinsieme C di A che abbia (x_0, y_0) come punto di accumulazione.

Sia $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ e sia (x_0, y_0) punto di accumulazione per A .

Se esistono due sottoinsiemi $C_1, C_2 \subseteq A$ tali che:

- (x_0, y_0) è punto di accumulazione per entrambi;

- esistono i limiti ristretti

$$\lim_{\substack{(x,y) \rightarrow (x_0,y_0) \\ (x,y) \in C_1}} f(x,y) = l_1$$

$$\lim_{\substack{(x,y) \rightarrow (x_0,y_0) \\ (x,y) \in C_2}} f(x,y) = l_2$$

- con $l_1 \neq l_2$,

allora

$$\lim_{(x,y) \rightarrow (x_0,y_0)} f(x,y) \quad \text{non esiste}$$

Se il limite totale esistesse ed fosse uguale a l , allora per ogni sottoinsieme $C \subseteq A$ tale che (x_0, y_0) sia punto di accumulazione per C , dovrebbe valere

$$\lim_{\substack{(x,y) \rightarrow (x_0,y_0) \\ (x,y) \in C}} f(x,y) = l$$

Questo contraddice l'ipotesi $l_1 \neq l_2$. Il risultato segue dall'unicità del limite.

Come condizione necessaria ma non sufficiente per l'esistenza del limite, si può verificare che i limiti lungo diverse curve (ad esempio rette $y = mx$, oppure curve come $y = x^2$) che tendono a (x_0, y_0) diano lo stesso valore. Se lungo due curve si ottengono due valori limite diversi, il limite della funzione non esiste.

3.2.2 Continuità

Definizione 3.2.5 (Continuità in un punto). Sia $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$. La funzione f si dice continua nel punto $P_0 = (x_0, y_0) \in A$ se

$$\lim_{(x,y) \rightarrow (x_0,y_0)} f(x,y) = f(x_0, y_0)$$

Definizione 3.2.6 (Continuità su un dominio). La funzione f si dice continua in A se è continua in ogni punto $P_0 \in A$.

Definizione 3.2.7 ($\epsilon - \delta$ per la continuità). La funzione f è continua in $P_0 = (x_0, y_0)$ se

$$\forall \varepsilon > 0, \exists \delta > 0 \text{ tale che}$$

$$\|(x, y) - (x_0, y_0)\| < \delta \Rightarrow |f(x, y) - f(x_0, y_0)| < \varepsilon.$$

DIMOSTRAZIONE

Sia $P_0 = (x_0, y_0)$.

Osserviamo che

$$|f(x, y) - f(x_0, y_0)| = |x - x_0|.$$

Poiché vale la diseguaglianza

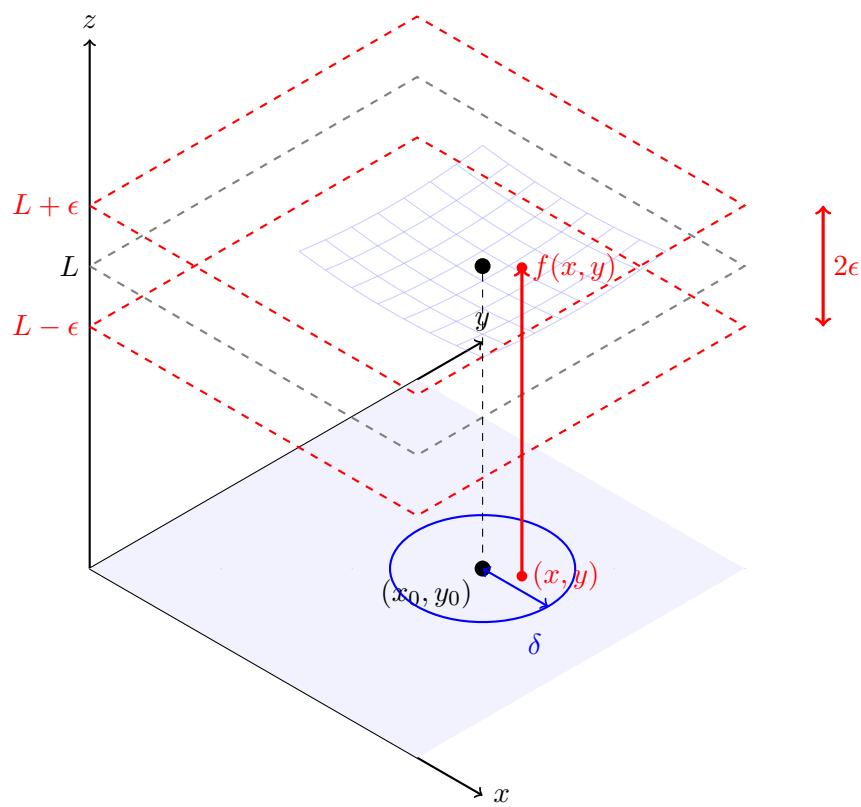
$$|x - x_0| \leq \sqrt{(x - x_0)^2 + (y - y_0)^2} = \|(x, y) - (x_0, y_0)\|,$$

se scegliamo $\delta = \varepsilon$, otteniamo:

$$\|(x, y) - (x_0, y_0)\| < \delta \Rightarrow |x - x_0| < \varepsilon.$$

Quindi f è continua in P_0 .

Poiché P_0 è arbitrario, f è continua in tutto \mathbb{R}^2 . □



Il grafico mostra la relazione tra "vicinanza nel dominio" e "vicinanza nel codominio".

Fissiamo una tolleranza ϵ sul valore della funzione (la banda rossa tra $L-\epsilon$ e $L+\epsilon$). La definizione di limite garantisce che esiste un raggio δ nel piano $x-y$ (il disco blu) tale che: se prendiamo qualunque punto (x, y) dentro questo disco, il corrispondente valore $f(x, y)$ sulla superficie cade necessariamente dentro la banda rossa.

Geometricamente, questo significa che la superficie $z = f(x, y)$ non presenta "salti" o discontinuità in (x_0, y_0) : restringendo progressivamente il disco blu (riducendo δ), i valori della funzione si concentrano sempre più attorno a L . Questo è l'essenza della continuità: piccole variazioni in (x, y) producono piccole variazioni in $f(x, y)$.

Teorema 3.2.1. *Siano f e g due funzioni a più variabili continue a valori reali, allora potremmo dire che*

- $f \pm g$ e $f \cdot g$ sono continue;
- $\frac{f}{g}$ sono continue se solo se $g \neq 0$;
- f^g è continua se solo se $g > 0$, perché corrisponde a $e^{g \cdot \ln(f)}$ dove $\ln(f)$ è continua per $f > 0$ che garantisce che sia ben definita, $g \cdot \ln(f)$ è continua per il prodotto di funzioni continue ed è continua;
- $g \circ f$ è continua dove è definita, ossia valore assunto da f deve appartenere al dominio di g .
- I polinomi sono sempre continui su tutto \mathbb{R}^n perché le funzioni coordinate x_1, x_2, \dots, x_n sono continue. Quindi, costruendo il polinomio a partire dalle coordinate tramite operazioni algebriche (somme, prodotti, multipli), otteniamo sempre una funzione continua.

Se abbiamo operazioni tra funzioni continue, il risultato è ancora continuo.

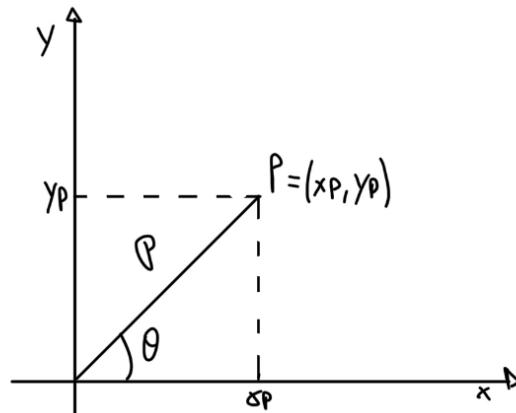
Questo teorema ci permette di affermare che le funzioni elementari, come polinomi a più variabili, funzioni razionali, funzioni esponenziali a più variabili, funzioni trigonometriche, sono tutte continue.

Alcuni esempi:

- $f(x, y) = \sin(x^2 + y^2)$ — continua su \mathbb{R}^2
- $f(x, y) = e^{xy}$ — continua su \mathbb{R}^2
- $f(x, y) = \ln(x^2 + y^2 + 1)$ — continua su \mathbb{R}^2 (argomento sempre positivo)
- $f(x, y) = \frac{x^2 - y^2}{x^2 + y^2 + 1}$ — continua su \mathbb{R}^2
- $f(x, y, z) = \sqrt{x^2 + y^2 + z^2}$ — continua su \mathbb{R}^3
- $f(x, y) = \arctan\left(\frac{y}{x}\right)$ — continua dove $x \neq 0$
- $f(x, y) = x^y$ con $x > 0$ — continua su $\{(x, y) : x > 0\}$

3.2.3 Coordinate Polari

Definizione 3.2.8 (Coordinate polari). Un modo alternativo alle coordinate cartesiane per descrivere posizioni sul piano sono le coordinate polari, un sistema bidimensionale che identifica un punto P attraverso due parametri: la distanza $\rho = \sqrt{(x - x_0)^2 + (y - y_0)^2}$ da un punto fisso O (detto polo), e l'angolo θ formato da una semiretta fissa uscente da O (detta asse polare) e il segmento che unisce O a P , misurato in senso antiorario.



Sappiamo bene il concetto di distanza (ρ), mentre dobbiamo spiegare bene il concetto di angolo (θ).

Geometricamente l'angolo θ è strettamente legato alla coefficiente angolare (o pendenza) della retta che passa per il punto P , ossia quanto ci alziamo verticalmente rispetto a quanto ci spostiamo orizzontalmente.

Dalla trigonometria sappiamo che il coefficiente angolare della retta è $\tan \theta = \frac{\sin \theta}{\cos \theta}$, che va a coincidere esattamente con la definizione di coefficiente angolare dell'angolo θ delle coordinate polari:

$$\tan \theta = \frac{\rho \sin \theta}{\rho \cos \theta} = \frac{y_\rho}{x_\rho}$$

Quindi se conosciamo le coordinate polari (x_ρ, y_ρ) possiamo trovare l'angolo θ invertendo la tangente e, sempre dalla trigonometria, sappiamo che $\tan^{-1} = \arctan$:

$$\theta = \arctan \frac{y_\rho}{x_\rho}$$

Per studiare il limite di una funzione $f(x, y)$ quando $(x, y) \rightarrow (x_0, y_0)$, si effettua una trasformazione in coordinate polari centrata nel punto limite.

Questo metodo è particolarmente efficace per funzioni radiali (che dipendono dalla distanza dal punto).

Procedimento:

1. Si pone il centro del sistema polare in (x_0, y_0) , il punto di accumulazione verso cui si tende, trasformando le coordinate cartesiane (x, y) seguendo il sistema

$$\begin{cases} x = x_0 + \rho \cos(\theta) \\ y = y_0 + \rho \sin(\theta) \end{cases}$$

2. Si studia il comportamento della nuova $f(\rho, \theta)$ per $\rho \rightarrow 0^+$, perché essendo una radice quadrata di una somma di quadrati, ρ non può mai essere negativa, quindi si avvicina al centro resitingendosi verso (x_0, y_0) ;
3. Si verifica che il limite sia uniforme rispetto a θ , ossia il valore non dipenda dalla direzione da cui ci si avvicina a (x_0, y_0) .

Il valore del limite non deve dipendere dalla direzione da cui ci si avvicina.

Definizione 3.2.9 ($\epsilon - \delta$).

$\forall \epsilon > 0 \exists \delta_\epsilon > 0$ tale che $\forall \rho$ con $0 < \rho < \delta$ e $\forall \theta \in [0, 2\pi]$ si ha $|f(x_0 + \rho \cos(\theta), y_0 + \rho \sin(\theta))| < \epsilon$

θ rappresenta tutte le possibili direzioni da cui ci si avvicina, se il limite dipende da θ significa che per cammini diversi otteniamo valori diversi e per l'esistenza del limite i valori devono essere gli stessi qualsiasi sia la direzione di avvicinamento.

Quindi, se il limite per $f(\rho, \theta) = g(\theta)$ per $\rho \rightarrow 0^+$ vuol dire che $g(\theta)$ assume valori diversi al variare di θ , violando l'uniformità rispetto a θ e affermando che il limite non esiste.

3.2.4 Il teorema del confronto per dimostrare l'esistenza di un limite

Definizione 3.2.10 (Il teorema del confronto (o dei due carabinieri)). Siano $f(x, y), g(x, y), h(x, y)$ tre funzioni con un intorno di $P_0 = (x_0, y_0)$. Sia

$$f(x, y) \leq g(x, y) \leq h(x, y)$$

Se il limite di $f(x, y)$ e $h(x, y)$ tende a l quando $(x, y) \rightarrow (x_0, y_0)$ allora il limite di $g(x, y)$ tende a l quando $(x, y) \rightarrow (x_0, y_0)$.

Questo perché la funzione centrale $g(x, y)$ è "intrappolata" tra le due, quindi il valore del suo limite dovrà per forza tendere a l .

Quindi, l'esistenza del limite è garantita se è possibile individuare una funzione maggiorante $g(\rho)$ tale che:

$$|f(x_0 + \rho \cos \theta, y_0 + \rho \sin \theta) - l| \leq g(\rho)$$

Affinché tale condizione sia valida, la funzione $g(\rho)$ deve possedere tre caratteristiche fondamentali:

- **Indipendenza da θ (Funzione Radiale):** g deve dipendere esclusivamente dalla distanza ρ . Questo garantisce che il valore del limite non dipenda dalla direzione scelta per avvicinarsi al punto (x_0, y_0) , assicurando l'uniformità della convergenza rispetto a θ ;
- **Non negatività:** Deve risultare $g(\rho) \geq 0$, in quanto deve maggiorare un valore assoluto (che rappresenta una distanza geometrica tra il valore della funzione e il suo limite);
- **Natura infinitesima:** Deve essere verificato che $\lim_{\rho \rightarrow 0^+} g(\rho) = 0$.

La necessità di una funzione $g(\rho)$ non negativa e infinitesima risiede nell'applicazione del teorema del confronto.

Sapendo che il valore assoluto di una differenza è per definizione non negativo, possiamo impostare la seguente catena di diseguaglianze:

$$0 \leq |f(x_0 + \rho \cos \theta, y_0 + \rho \sin \theta) - l| \leq g(\rho)$$

Se la funzione maggiorante $g(\rho)$ tende a zero per $\rho \rightarrow 0^+$, allora anche il termine centrale è costretto a tendere a zero. Formalmente:

1. Si maggiora lo scarto: $|f(\rho, \theta) - \ell| \leq g(\rho)$.
2. Si verifica il comportamento al limite $\lim_{\rho \rightarrow 0^+} g(\rho) = 0$.
3. Per il Teorema dei Carabinieri, si conclude che:

$$\lim_{\rho \rightarrow 0^+} |f(x_0 + \rho \cos \theta, y_0 + \rho \sin \theta) - \ell| = 0$$

Tale risultato implica univocamente che $f \rightarrow \ell$, confermando l'esistenza del limite indipendentemente dalla direzione θ con cui ci si avvicina al punto (x_0, y_0) .

Studio della continuità

Sia $f : A \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$ e sia un punto $P = (x_0, y_0)$, allora f è continua in P se rispetta le condizioni:

- f sia definita in $P = (x_0, y_0) \in D$ dominio di f ;
- Esiste il limite $\lim_{(x,y) \rightarrow (x_0,y_0)} f(x, y) = \ell$ e $\ell = f(x_0, y_0)$

Il procedimento meccanico per dimostrare la continuità che coinvolge anche le coordinate polari:

- $P = (x_0, y_0) \in D$ sia definita nel dominio di f ;
- Si trasforma le coordinate cartesiane in coordinate polari, si controlla se esiste il limite $\lim_{\rho \rightarrow 0^+} f(\rho, \theta) = 0$ e con le dovute analisi di maggiorazione si controlla che il limite dipenda uniformemente rispetto a θ verificando l'uguaglianza.

Calcolo Differenziale

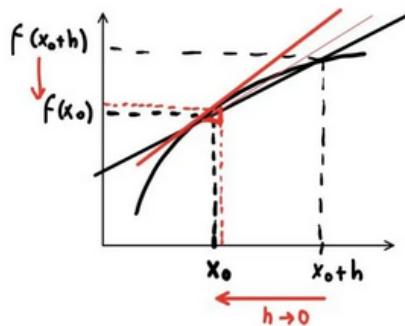
4.1 Derivata

4.1.1 Derivata parziale

Se $f : I \subseteq \mathbb{R} \rightarrow \mathbb{R}$ si ha che la derivata $f'(x)$ o $\frac{df(x)}{dx}$ rappresenta il tasso di variazione di f in un punto. Matematicamente $f'(x)$ è definito come, se esiste finito, il limite

$$\lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}, \quad \forall x \in \mathbb{R}$$

Quindi, $f'(x)$ è il coefficiente angolare della retta tangente al grafico di $f(x)$ nel punto $P = (x, f(x))$.



Definizione 4.1.1 (Derivata parziale rispetto ad una variabile). *La derivata parziale di una funzione $f : \mathbb{R}^n \rightarrow \mathbb{R}$ rispetto ad una variabile k è la derivata ordinaria di f considerata come funzione delle sola variabile k , mantenendo "fisse" le altre variabili.*

Il termine "parziale" indica che si sta considerando solo una parte del dominio, ovvero la variazione della funzione rispetto a una singola variabile, "congelando" tutte le altre. È una derivata parziale nel senso che cattura solo un aspetto del comportamento della funzione, non la sua variazione completa in tutte le direzioni.

La derivata parziale fornisce informazioni locali lungo una direzione specifica (quella dell'asse coordinata corrispondente), ma non descrive il comportamento globale della funzione in un intorno del punto.

Formalmente, sia un punto $P_0 = (x_0, y_0) \in A$ e sia $f : A \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$ con A aperto:

- La derivata rispetto a x (solo x incrementata)

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

- La derivata rispetto a y (solo y incrementata)

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

Se questi due limiti esistono finiti, si dirà che f è derivabile parzialmente (o localmente) rispetto a x e a y usando la notazione $\frac{\partial f}{\partial x_0}(x_0, y_0)$ e $\frac{\partial f}{\partial y_0}(x_0, y_0)$.

Potremmo dire che la derivata parziale è una derivata ordinaria di f sulle tracce parallele agli assi delle coordinate, lungo il piano o $x = x_0$ o $y = y_0$:

- $\frac{\partial f}{\partial x}(x_0, y_0) = g'(x_0) = \frac{d}{dx}f(x, y_0)$
- $\frac{\partial f}{\partial y}(x_0, y_0) = h'(y_0) = \frac{d}{dy}f(x_0, y)$

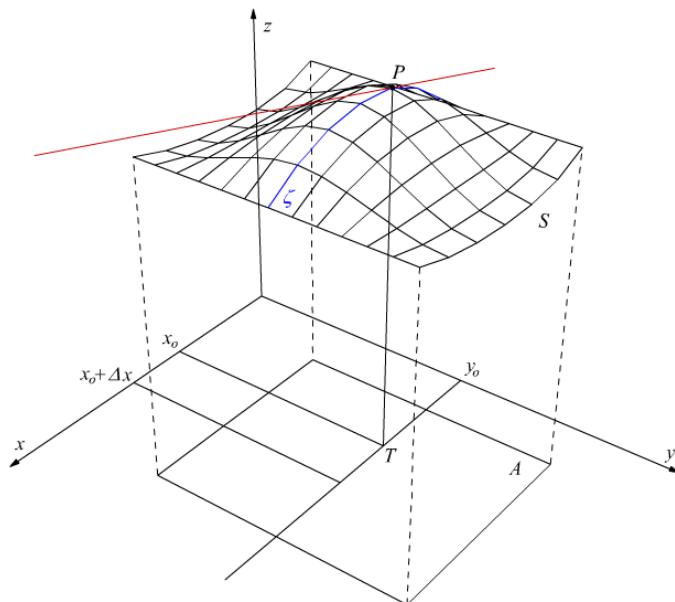
Definizione 4.1.2 (Derivata parziale rispetto a n variabili). Sia $f : A \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$ con A aperto e sia un punto $P_0 = (\underline{x}) \in A$. La derivata parziale della f rispetto alla variabile x_i con $i = 1 \dots n$ in $P_0 = (\underline{x}) = (x_1 \dots x_n)$ corrisponde a

$$\lim_{h \rightarrow 0} \frac{f(x_1, x_2, \dots, x_{i-1}, x_i + h, \dots, x_n) - f(x_1, \dots, x_n)}{h}$$

Se tale limite esiste ed è finito, esiste la derivata parziale e si scrive come:

$$\frac{\partial f}{\partial x_i}(\underline{x}) = f_{x_i}(\underline{x}) = D_{x_i}f(\underline{x})$$

Geometricamente, nelle funzioni a due variabili le derivate parziali sono legate al concetto di rette tangenti al grafico delle curve ottenute fissata la traccia xz o yz .



4.1.2 Gradiente

Una funzione $f : A \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$ con A aperto è derivabile su un punto $P_0 = (x_0, y_0) \in A$ se esistono le derivate parziali $\frac{\partial f}{\partial x_0}(x_0, y_0)$ e $\frac{\partial f}{\partial y_0}(x_0, y_0)$, o meglio, se esiste il vettore gradiente

$$\nabla f(x_0, y_0) = \left(\frac{\partial f}{\partial x}(x_0, y_0), \frac{\partial f}{\partial y}(x_0, y_0) \right)$$

Nel caso generale di una funzione $f : \mathbb{R}^n \rightarrow \mathbb{R}$:

$$\nabla f(\underline{x}) = \left(\frac{\partial f}{\partial x_1}(\underline{x}), \dots, \frac{\partial f}{\partial x_n}(\underline{x}) \right)$$

Affinché una funzione multivariata sia derivabile in un punto:

- f deve essere continua nel punto;
- Devono esistere tutte le derivate nel punto, quindi deve esistere il gradiente tale che $\nabla f(\underline{x}) = (\frac{\partial f}{\partial x_1}(\underline{x}), \dots, \frac{\partial f}{\partial x_n}(\underline{x}))$.

Avendo più direzioni lungo cui la funzione può variare, dobbiamo "memorizzare" n tassi di variazione. Di conseguenza, le n derivate parziali vengono organizzate in un vettore n -dimensionale che contiene tutta l'informazione sulla variazione locale della funzione.

Definizione 4.1.3 (Gradiente). Il gradiente è un vettore denotato $\nabla f \in \mathbb{R}^n$ le cui componenti sono le derivate parziali della funzione f nel punto \underline{x} :

$$\nabla f(\underline{x}) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(\underline{x}) \\ \frac{\partial f}{\partial x_2}(\underline{x}) \\ \vdots \\ \frac{\partial f}{\partial x_n}(\underline{x}) \end{pmatrix} \in \mathbb{R}^n$$

Esso raccoglie tutte le derivate parziali e misura quanto f è "sensibile" a variazioni in ciascuna variabile.

Da qui in poi non abbiamo più una retta tangente ma un piano tangente, ovvero un'approssimazione lineare tramite il prodotto scalare:

$$f(\underline{x} + \underline{h}) \approx f(\underline{x}) + \langle \nabla f(\underline{x}), \underline{h} \rangle + o(\|\underline{h}\|)$$

- **In una variabile:** vicino a un punto x_0 vale l'approssimazione lineare

$$f(x_0 + h) \approx f(x_0) + f'(x_0)h$$

dove h è una piccola variazione di x , $f'(x_0)$ è la pendenza della tangente, e $f'(x_0)h$ rappresenta di quanto cambia f quando x cambia di h .

- **In più variabili:** il punto non è più un numero ma un vettore $\underline{x} = (x_1, x_2, \dots, x_n)$ e anche la variazione non è più un numero ma un vettore $\underline{h} = (h_1, h_2, \dots, h_n)$.

Si usa il prodotto scalare perché quando ti sposti da \underline{x} a $\underline{x} + \underline{h}$ stai cambiando tutte le variabili della funzione insieme, ossia sommando il contributo di ogni variabile moltiplicato per la corrispondente derivata parziale:

$$\langle \nabla f(\underline{x}), \underline{h} \rangle = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\underline{x}) \cdot h_i$$

Quindi,

$$f(\underline{x} + \underline{h}) \approx f(\underline{x}) + \langle \nabla f(\underline{x}), \underline{h} \rangle$$

Vicino a \underline{x} , la funzione si comporta come un piano tangente: il termine con il prodotto scalare fornisce l'approssimazione lineare migliore e l'errore diventa trascurabile quando $\|\underline{h}\|$ è piccolo.

4.1.3 Derivata direzionale

Sia $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ e siano $\underline{x} = \{x_1, \dots, x_n\} \in A$ e e_i versore componenente della base canonica $e = (e_1, \dots, e_n)$ di \mathbb{R}^n (direzione nello spazio).

La derivata direzionale rispetto a f rispetto alla direzione di e_i nel punto \underline{x} se esiste allora esiste finito

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

- $he_i = h(0\dots 1\dots 0) = (0\dots h\dots 0)$
- $\underline{x} + he_i = (x_1\dots x_i + h\dots x_n)$
- f dipende da h

La derivata parziale può essere rappresentata con un concetto più generale: la derivata direzionale. Dato un piano tangente, per il punto in cui si poggia il piano passano infinite rette.

Informalmente, la derivata direzionale è la pendenza di una di quelle rette in riferimento ad una direzione scelta partendo dal punto del dominio.

Definizione 4.1.4 (Derivata direzionale). *Dato un versore \underline{v} , ovvero un vettore di norma unitaria ($\|\underline{v}\| = 1$), in \mathbb{R}^n , sia $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ con A aperto e sia $\underline{x} = (x_1\dots x_n) \in A$.*

La derivata direzionale di f rispetto alla direzione data da \underline{v} nel punto \underline{x} esiste se esiste ed è finito il limite

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

L'uso del versore risulta importante perché non interessa la lunghezza del vettore ma solamente la direzione e il verso, quindi se non è di norma unitaria va normalizzato con la formula $v = \frac{\underline{v}}{\|\underline{v}\|}$.

4.2 Differenziabilità

Differenza tra derivabilità e differenziabilità

- **Analisi 1 (Caso unidimensionale):**

Per le funzioni di una variabile, non c'è differenza tra derivabilità e differenziabilità.

La derivabilità garantisce l'esistenza della retta tangente al grafico della funzione in un punto x_0 .

La differenziabilità garantisce che la funzione può essere sostituita localmente con la retta tangente.

In modo più preciso, se f è differenziabile in x_0 , allora:

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + o(x - x_0)$$

dove il termine $o(x - x_0)$ indica una quantità che tende a zero più velocemente di $x - x_0$.

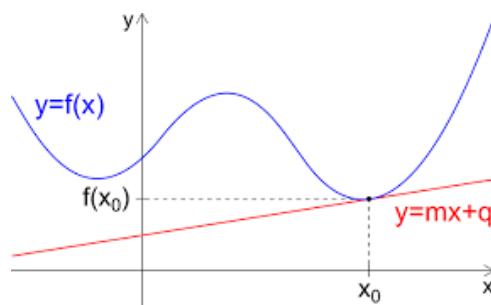
Se dividiamo entrambi i membri per $x - x_0$ e prendiamo il limite per $x \rightarrow x_0$, otteniamo:

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = f'(x_0)$$

Questo significa che:

$$f(x) - \text{retta tangente} = o(x - x_0)$$

e quindi, avvicinandosi a x_0 , la funzione e la sua approssimazione lineare diventano sempre più indistinguibili, indipendentemente da quanto piccolo sia l'intorno scelto.



• **Analisi 2 (Caso multivariato):**

In più variabili ($f : \mathbb{R}^n \rightarrow \mathbb{R}$) i concetti di derivabilità (esistenza delle derivate parziali) e differenziabilità si separano.

Una funzione $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ è differenziabile in \underline{x}_0 se l'incremento della funzione può essere approssimato linearmente in modo globale.

Esiste un'applicazione lineare $L : \mathbb{R}^n \rightarrow \mathbb{R}$ tale che:

$$f(\underline{x}_0 + \underline{h}) = f(\underline{x}_0) + L(\underline{h}) + o(\|\underline{h}\|) \quad \text{per } \underline{h} \rightarrow \underline{0}$$

dove

- $L(\underline{h}) = \langle \nabla f(\underline{x}_0), \underline{h} \rangle$ è l'approssimazione lineare (il piano tangente)
- $o(\|\underline{h}\|)$ è l'errore, che tende a zero più velocemente di $\|\underline{h}\|$

Se f è differenziabile, l'applicazione lineare L è univocamente determinata dal gradiente: $L(\underline{h}) = \nabla f(\underline{x}_0) \cdot \underline{h}$.

Esplicitamente in due variabili, ponendo $\underline{h} = (x - x_0, y - y_0)$, la formula diventa:

$$f(x, y) = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) + o\left(\sqrt{(x - x_0)^2 + (y - y_0)^2}\right)$$

La parola globale significa che questa approssimazione funziona in tutte le direzioni contemporaneamente: non importa come ci avviciniamo a \underline{x}_0 (lungo quale curva o direzione), l'errore relativo:

$$\lim_{\underline{h} \rightarrow \underline{0}} \frac{|f(\underline{x}_0 + \underline{h}) - f(\underline{x}_0) - L(\underline{h})|}{\|\underline{h}\|} = 0$$

tende sempre a zero, indipendentemente dalla direzione scelta.

In contrasto, l'esistenza delle derivate parziali garantisce solo approssimazioni lineari lungo gli assi coordinati (direzioni "locali"). La differenziabilità richiede che l'approssimazione funzioni uniformemente in tutte le direzioni.

Esempio fondamentale (Derivabile ma non Differenziabile):

Consideriamo la funzione

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

Nel punto $(0, 0)$:

1. **Le derivate parziali esistono e sono nulle:** infatti, lungo gli assi $x = 0$ e $y = 0$ la funzione è identicamente nulla, quindi $f_x(0, 0) = 0$ e $f_y(0, 0) = 0$. Il candidato piano tangente sarebbe il piano $z = 0$;
2. **La funzione non è differenziabile:** se proviamo ad avvicinarci all'origine lungo la bisettrice $y = x$, otteniamo $f(x, x) = x/2$. La pendenza è $1/2$, mentre il gradiente nullo prevedeva pendenza 0. L'approssimazione lineare fallisce.

Questo dimostra che possiamo avere le derivate parziali (derivabilità) senza avere un piano tangente che approssima bene la funzione in tutte le direzioni (differenziabilità).

Per evitare di verificare il limite ogni volta, si usa il seguente teorema operativo:

Teorema 4.2.1 (Condizione sufficiente per la differenziabilità). *Se tutte le derivate parziali di f esistono e sono continue in un intorno di \underline{x}_0 (ovvero $f \in C^1$), allora f è differenziabile in \underline{x}_0 .*

La continuità delle derivate "stabilizza" le pendenze, impedendo comportamenti patologici e garantendo l'esistenza del piano tangente.

4.2.1 Il concetto di funzione differenziabile e l'esistenza del piano tangente

Definizione 4.2.1 (Differenziabilità). Una funzione $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ è differenziabile in $\underline{x}_0 \in A$ se e solo se:

- f è derivabile in \underline{x}_0 , ovvero $\exists \nabla f(\underline{x}_0)$;
- Vale il seguente limite:

$$\lim_{\underline{h} \rightarrow \underline{0}} \frac{f(\underline{x}_0 + \underline{h}) - f(\underline{x}_0) - \langle \nabla f(\underline{x}_0) \cdot \underline{h} \rangle}{\|\underline{h}\|} = 0$$

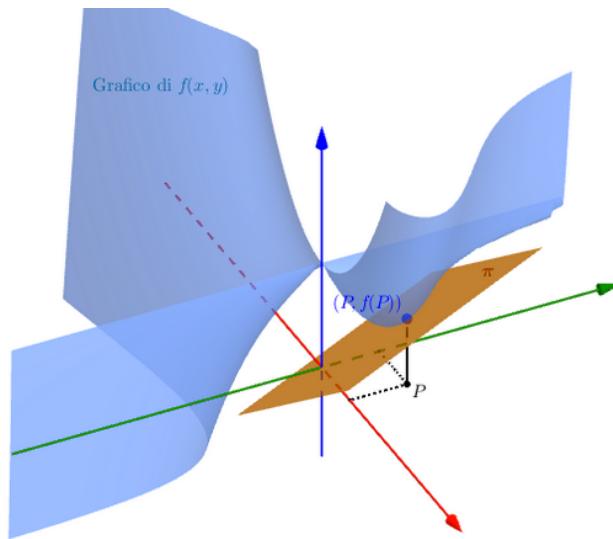
Dove $\underline{h} = (h_1, \dots, h_n)$ è il vettore incremento e $\|\underline{h}\| = \sqrt{\sum h_i^2}$ è la sua norma euclidea.

Affermare che il limite è 0 significa che il numeratore (l'errore dell'approssimazione lineare) tende a zero con un ordine di infinitesimo superiore rispetto al denominatore (la distanza percorsa $\|\underline{h}\|$). In simboli:

$$f(\underline{x}_0 + \underline{h}) - f(\underline{x}_0) - \nabla f(\underline{x}_0) \cdot \underline{h} = o(\|\underline{h}\|)$$

Nota. Geometricamente, dire che una funzione $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ è differenziabile in un punto $P_0 = (x_0, y_0)$ equivale a dire che il grafico della funzione ammetta un piano tangente non verticale in $P_0 = (\underline{x}_0, f(\underline{x}_0))$.

Non basta che il piano tocchi il punto (quello lo fanno infiniti piani). Significa che il piano è talmente "aderente" alla superficie curva che, se "zoomiamo" infinitamente sul punto P_0 , la superficie curva del grafico diventa indistinguibile dal piano piatto. L'errore che commettiamo sostituendo la superficie con il piano tende a zero più velocemente della distanza dal punto.



Equazione del piano tangente

Sia $f : A \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$ differenziabile in $P_0 = (x_0, y_0)$. L'equazione del piano tangente al grafico in P_0 è:

$$z = f(x_0, y_0) + \langle \nabla f(P_0) \cdot \underline{h} \rangle$$

Dove $\underline{h} = (x - x_0, y - y_0)$ è il vettore spostamento. Esplicitando in componenti:

$$z = \underbrace{f(x_0, y_0)}_{\text{punto di contatto}} + \underbrace{\frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0)}_{\text{inclinazione}}$$

Questa formula è la naturale estensione geometrica della retta tangente. In dimensione 1, l'equazione della retta tangente è:

$$y = f(x_0) + f'(x_0)(x - x_0)$$

Mentre l'approssimazione della funzione (formula di Taylor al primo ordine) è:

$$f(x) = \underbrace{f(x_0) + f'(x_0)(x - x_0)}_{\text{Parte Lineare (Retta)}} + \underbrace{o(x - x_0)}_{\text{Errore}} \quad \text{per } x \rightarrow x_0$$

Allo stesso modo, in \mathbb{R}^2 , il piano tangente rappresenta la parte lineare dello sviluppo della funzione.

4.2.2 Continuità di una funzione differenziabile e il teorema del differenziale totale

Teorema 4.2.2 (Continuità di una funzione differenziabile). *Sia $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ differenziabile in un punto $\underline{x}_0 \in A$ allora è continua in \underline{x}_0 .*

DIMOSTRAZIONE

Per dimostrare la continuità in \underline{x}_0 , dobbiamo verificare che:

$$\lim_{\underline{h} \rightarrow \underline{0}} f(\underline{x}_0 + \underline{h}) = f(\underline{x}_0)$$

Dalla definizione di differenziabilità, possiamo scrivere l'incremento della funzione come somma di una parte lineare e un infinitesimo di ordine superiore:

$$f(\underline{x}_0 + \underline{h}) = f(\underline{x}_0) + \langle \nabla f(\underline{x}_0) \cdot \underline{h} \rangle + o(\|\underline{h}\|) \quad \text{per } \underline{h} \rightarrow \underline{0}$$

Passando al limite per $\underline{h} \rightarrow \underline{0}$ in entrambi i membri:

$$\lim_{\underline{h} \rightarrow \underline{0}} f(\underline{x}_0 + \underline{h}) = \lim_{\underline{h} \rightarrow \underline{0}} [f(\underline{x}_0) + \langle \nabla f(\underline{x}_0) \cdot \underline{h} \rangle + o(\|\underline{h}\|)]$$

Sfruttando la linearità del limite (il limite della somma è la somma dei limiti), analizziamo i tre termini separatamente:

- Il primo termine è costante rispetto ad \underline{h} (non dipende dall'incremento), quindi:

$$\lim_{\underline{h} \rightarrow \underline{0}} f(\underline{x}_0) = f(\underline{x}_0)$$

- Il termine lineare (prodotto scalare) tende a 0. Possiamo dimostrarlo usando la disegualanza di Cauchy-Schwarz e il teorema del confronto. Notiamo che:

$$0 \leq |\nabla f(\underline{x}_0) \cdot \underline{h}| \leq \|\nabla f(\underline{x}_0)\| \cdot \|\underline{h}\|$$

Poiché $\|\nabla f(\underline{x}_0)\|$ è un numero finito e $\|\underline{h}\| \rightarrow 0$, il prodotto tende a zero. Di conseguenza:

$$\lim_{\underline{h} \rightarrow \underline{0}} (\nabla f(\underline{x}_0) \cdot \underline{h}) = 0$$

- L'ultimo termine tende a zero per la definizione stessa di o-piccolo (errore infinitesimo):

$$\lim_{\underline{h} \rightarrow \underline{0}} o(\|\underline{h}\|) = 0$$

Mettendo tutto insieme otteniamo:

$$\lim_{\underline{h} \rightarrow \underline{0}} f(\underline{x}_0 + \underline{h}) = f(\underline{x}_0) + 0 + 0 = f(\underline{x}_0)$$

Poiché il limite coincide con il valore della funzione nel punto, f è continua in \underline{x}_0 . □

Il differenziale

Nell'analisi a più variabili, dato un punto base \underline{x}_0 e un vettore spostamento \underline{h} :

- **L'incremento vero (Δf):** Rappresenta la variazione reale della funzione muovendosi sulla superficie curva del grafico.

$$\Delta f = f(\underline{x}_0 + \underline{h}) - f(\underline{x}_0)$$

- **Il differenziale (df):** Rappresenta la variazione calcolata muovendosi sul piano tangente. È la parte lineare dell'incremento.

$$df = \nabla f(\underline{x}_0) \cdot \underline{h}$$

Mentre Δf misura il dislivello esatto sulla "montagna" (il grafico di f), il differenziale df misura il dislivello che avremmo se camminassimo sul piano tangente che approssima la montagna in quel punto.

Per spostamenti \underline{h} molto piccoli (in un intorno di \underline{x}_0), l'errore commesso sostituendo la curva con il piano è trascurabile, quindi:

$$\Delta f \approx df$$

Più formalmente, la differenza $\Delta f - df$ è un infinitesimo di ordine superiore rispetto alla norma dello spostamento ($o(\|\underline{h}\|)$).

Teorema 4.2.3 (Teorema del differenziale totale). *Sia $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ con A aperto e sia f di classe C^1 su A , ovvero tutte le derivate parziali esistono e sono continue in un punto $\underline{x} \in A$:*

$$\exists \nabla f(\underline{x}) \Leftrightarrow \text{esistono } f_{x_1}(\underline{x}), f_{x_2}(\underline{x}), \dots, f_{x_n}(\underline{x})$$

Allora f è differenziabile in \underline{x} .

DIMOSTRAZIONE su $n = 2$ (vale per n variabili)

Sia $f : A \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$ tale che $(x, y) \rightarrow f(x, y)$, si suppone che le derivate rispetto alle due variabili esistano $\forall (x, y) \in A$ esistano e siano continue, allora mostro che f è differenziabile nel punto $(x, y), \forall \underline{h} \in \mathbb{R}^2$ con $\underline{h} = (h, k)$

$$\lim_{h,k \rightarrow 0} \frac{f(x+h, y+k) - f(x, y) - \langle \nabla f(x, y) \cdot (h, k) \rangle}{\sqrt{h^2 + k^2}} = 0$$

Passo 1: Consideriamo l'incremento della funzione:

$$\Delta f = f(x+h, y+k) - f(x, y)$$

Aggiungiamo e sottraiamo il termine misto $f(x, y+k)$

$$= \underbrace{[f(x+h, y+k) - f(x, y+k)]}_{\text{Variazione solo in } x} + \underbrace{[f(x, y+k) - f(x, y)]}_{\text{Variazione solo in } y}$$

Perché nel primo membro la y in $y+k$ rimane costante, nel secondo membro invece è la x ad essere costante.

Passo 2: Applicazione del Teorema del valor medio (Lagrange)

Poiché f è derivabile, applichiamo il Teorema del valor medio separatamente alle due parentesi:

- Per la prima parentesi (variabile x), esiste un punto x_1 compreso tra x e $x+h$ tale che:

$$f(x+h, y+k) - f(x, y+k) = f_x(x_1, y+k) \cdot (x+h - x) = f_x(x_1, y+k) \cdot h \text{ con } h > 0$$

- Per la seconda parentesi (variabile y), esiste un punto y_1 compreso tra y e $y + k$ tale che:

$$f(x, y + k) - f(x, y) = f_y(x, y_1) \cdot (y + k - y) = f_y(x, y_1) \cdot k \text{ con } k > 0$$

Quindi l'incremento diventa:

$$\Delta f = f_x(x_1, y + k)h + f_y(x, y_1)k$$

Passo 3: Costruzione del limite di differenziabilità

Dobbiamo verificare che:

$$\lim_{(h,k) \rightarrow (0,0)} \frac{f(x+h, y+k) - f(x, y) - [f_x(x, y)h + f_y(x, y)k]}{\sqrt{h^2 + k^2}} = 0$$

Sostituiamo Δf con l'espressione trovata con il teorema del valor medio. Il numeratore diventa:

$$\text{Num} = [f_x(x_1, y + k)h + f_y(x, y_1)k] - [f_x(x, y)h + f_y(x, y)k]$$

Raccogliamo h e k :

$$\text{Num} = [f_x(x_1, y + k) - f_x(x, y)]h + [f_y(x, y_1) - f_y(x, y)]k$$

Passo 4: Maggiorazione con il valore assoluto

Passiamo al valore assoluto e usiamo la disegualanza triangolare ($|a + b| \leq |a| + |b|$):

$$\left| \frac{\text{Num}}{\sqrt{h^2 + k^2}} \right| \leq \frac{|f_x(x_1, y + k) - f_x(x, y)| \cdot |h|}{\sqrt{h^2 + k^2}} + \frac{|f_y(x, y_1) - f_y(x, y)| \cdot |k|}{\sqrt{h^2 + k^2}}$$

Osserviamo che i rapporti $\frac{|h|}{\sqrt{h^2 + k^2}}$ e $\frac{|k|}{\sqrt{h^2 + k^2}}$ sono sempre ≤ 1 (poiché il cateto è minore dell'ipotenusa). Quindi:

$$0 \leq \text{Limite} \leq |f_x(x_1, y + k) - f_x(x, y)| + |f_y(x, y_1) - f_y(x, y)|$$

Passo 5: Conclusione per continuità

Facciamo tendere $(h, k) \rightarrow (0, 0)$.

- Se $h \rightarrow 0$, allora $x_1 \rightarrow x$ (perché compreso tra x e $x + h$).
- Se $k \rightarrow 0$, allora $y_1 \rightarrow y$.

Poiché per ipotesi f è di classe C^1 , le derivate parziali f_x e f_y sono continue. Pertanto:

$$\lim_{(h,k) \rightarrow (0,0)} f_x(x_1, y + k) = f_x(x, y)$$

$$\lim_{(h,k) \rightarrow (0,0)} |f_x(x_1, y + k) - f_x(x, y)| = 0$$

E

$$\lim_{(h,k) \rightarrow (0,0)} f_y(x, y_1) = f_y(x, y)$$

$$\lim_{(h,k) \rightarrow (0,0)} |f_y(x, y_1) - f_y(x, y)| = 0$$

Essendo dei valori assoluti tendono a zero quando $(h, k) \rightarrow (0, 0)$, quindi si può affermare che il limite è nullo, allora la funzione è differenziabile.

□

Gerarchia di regolarità

Sia $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ con A aperto, definiamo $C^1(A)$ come l'insieme delle funzioni con derivate parziali continue su A .

Vale la seguente catena di implicazioni per funzioni di più variabili:

$$f \in C^1(A) \implies f \text{ differenziabile in } A \implies f \in C^0(A) \text{ continua in ogni punto di } A$$

$f \in C^1$ è una condizione *sufficiente* ma non necessaria per la differenziabilità. Esistono funzioni differenziabili che hanno derivate parziali discontinue (casi rari e patologici).

Relazione tra differenziabilità e derivate direzionali

Definizione 4.2.2 (La differenziabilità implica la continuità e l'esistenza delle derivate direzionali). Sia $f : A \rightarrow \mathbb{R}$, con $A \subset \mathbb{R}^n$ aperto e $x_0 \in A$. Se f è differenziabile in \underline{x}_0 con differenziale L , allora:

1. f è continua in \underline{x}_0 ;
2. f ammette tutte le derivate direzionali in \underline{x}_0 e

$$D_{\vec{v}}f(\underline{x}_0) = L(\vec{v}), \quad \forall \vec{v} \in \mathbb{R}^n.$$

DIMOSTRAZIONE

In particolare, si può dimostrare che il differenziale L , se esiste, è unico.

Sia $\underline{v} \in \mathbb{R}^n$, per definizione, la derivata direzionale è:

$$D_{\underline{v}}f(\underline{x}_0) = \lim_{t \rightarrow 0} \frac{f(\underline{x}_0 + t\underline{v}) - f(\underline{x}_0)}{t}$$

Usando la definizione di differenziabilità con incremento $\underline{h} = t\underline{v}$, si ottiene:

$$f(\underline{x}_0 + t\underline{v}) = f(\underline{x}_0) + L(t\underline{v}) + o(\|t\underline{v}\|)$$

Per la linearità di L , sappiamo che $L(t\underline{v}) = tL(\underline{v})$. Sostituendo e portando $f(\underline{x}_0)$ a sinistra:

$$f(\underline{x}_0 + t\underline{v}) - f(\underline{x}_0) = tL(\underline{v}) + o(|t|\|\underline{v}\|)$$

Dividendo tutto per $t \neq 0$:

$$\frac{f(\underline{x}_0 + t\underline{v}) - f(\underline{x}_0)}{t} = L(\underline{v}) + \frac{o(|t|\|\underline{v}\|)}{t}$$

Passando al limite per $t \rightarrow 0$, il termine con l'o-piccolo tende a 0. Segue che:

$$D_{\underline{v}}f(\underline{x}_0) = L(\underline{v})$$

Poiché il limite (la derivata direzionale) è unico, anche l'applicazione lineare L deve essere unica. □

Definizione 4.2.3 (La formula del gradiente). Sia f differenziabile su A e sia un vettore \vec{v} unitario, si può definire la derivata direzionale con la formula

$$D_{\vec{v}}f(x, y) = \langle \nabla f(x, y), \vec{v} \rangle$$

4.2.3 La regola della catena

Sia $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ una funzione differenziabile su un aperto A .

Sia $\gamma : I \subseteq \mathbb{R} \rightarrow A \subseteq \mathbb{R}^n$ una curva regolare (derivabile con derivata continua).

Allora la funzione composta

$$F(t) = F((f \circ \gamma)(t)) = f(\gamma(t))$$

è derivabile in ogni punto $t \in I$ e la sua derivata la derivata equivale al prodotto scalare tra il gradiente di f e il vettore tangente alla curva γ :

$$F'(t) = \langle \nabla f(\gamma(t)), \gamma'(t) \rangle$$

nel caso $n = 2$ con $\gamma(t) = (x(t), y(t))$ si ha

$$F'(t) = f_x(x(t), y(t)) \cdot x'(t) + f_y(x(t), y(t)) \cdot y'(t)$$

DIMOSTRAZIONE

Passo 1: Consideriamo il rapporto incrementale di $F(t)$ in un punto $t \in I$ rispetto a un incremento $h \in \mathbb{R}$.

Poiché per ipotesi f è differenziabile nel punto $\underline{x} = \gamma(t)$, per definizione di differenziabilità possiamo scrivere l'incremento di f come somma della parte lineare (gradiente) e dell'errore (o-piccolo). Allora:

$$f(\gamma(t+h)) - f(\gamma(t)) = \langle \nabla f(\gamma(t)), \gamma(t+h) - \gamma(t) \rangle + o(\|\gamma(t+h) - \gamma(t)\|)$$

Sostituiamo questa espressione nel rapporto incrementale:

$$\frac{F(t+h) - F(t)}{h} = \frac{\langle \nabla f(\gamma(t)), \gamma(t+h) - \gamma(t) \rangle}{h} + \frac{o(\|\gamma(t+h) - \gamma(t)\|)}{h}$$

Per la linearità del prodotto scalare, possiamo portare $1/h$ dentro il primo termine:

$$\frac{F(t+h) - F(t)}{h} = \left\langle \nabla f(\gamma(t)), \frac{\gamma(t+h) - \gamma(t)}{h} \right\rangle + \frac{o(\|\gamma(t+h) - \gamma(t)\|)}{h}$$

Passo 2: Passiamo ora al limite per $h \rightarrow 0$.

Poiché γ è derivabile in t , il rapporto incrementale vettoriale tende alla derivata $\gamma'(t)$:

$$\lim_{h \rightarrow 0} \frac{\gamma(t+h) - \gamma(t)}{h} = \gamma'(t)$$

Quindi il primo termine tende a:

$$\langle \nabla f(\gamma(t)), \gamma'(t) \rangle$$

Che è esattamente la tesi che cerchiamo.

Passo 3: Dobbiamo ora dimostrare che il resto tende a 0.

$$\lim_{h \rightarrow 0} \frac{o(\|\gamma(t+h) - \gamma(t)\|)}{h} = 0$$

Supponendo sia non nullo, riscriviamo

$$\left| \frac{o(\|\gamma(t+h) - \gamma(t)\|)}{h} \right|$$

Moltiplichiamo e dividiamo per la norma dell'incremento $\|\gamma(t+h) - \gamma(t)\|$:

$$\lim_{h \rightarrow 0} \underbrace{\frac{o(\|\gamma(t+h) - \gamma(t)\|)}{\|\gamma(t+h) - \gamma(t)\|}}_{(A)} \cdot \underbrace{\frac{\|\gamma(t+h) - \gamma(t)\|}{|h|}}_{(B)}$$

Analizziamo i fattori:

- **Fattore (A):** Tende a 0 per definizione di o-piccolo. Infatti, per la continuità di γ , quando $h \rightarrow 0$ anche l'incremento $\|\gamma(t+h) - \gamma(t)\| \rightarrow 0$.
- **Fattore (B):**

$$\lim_{h \rightarrow 0} \left[\sum_{i=1}^n \left(\frac{\gamma_i(t+h) - \gamma_i(t)}{h} \right)^2 \right]^{\frac{1}{2}} = \left[\sum_{i=1}^n (\gamma'_i(t))^2 \right]^{\frac{1}{2}} \in \mathbb{R}$$

Tende alla norma della derivata $\|\gamma'(t)\|$, che è un numero reale finito (la derivata esiste per ipotesi)

Quindi abbiamo un termine che tende a 0 moltiplicato per una quantità limitata ($\|\gamma'(t)\|$). Il limite complessivo è 0.

Ricomponendo i pezzi, il limite del rapporto incrementale è:

$$F'(t) = \langle \nabla f(\gamma(t)), \gamma'(t) \rangle + 0$$

□

Applicando la regola della catena possiamo rispondere alla domanda: in quale direzione si ha la massima crescita di una funzione differenziabile?

L'idea fondamentale è trasformare un problema in più variabili in un problema in una sola variabile, dove il concetto di pendenza è immediato.

Passo 1: Il punto e il movimento

Sia una funzione differenziabile $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ e sia $\underline{x}_0 \in A$ un punto fissato.

Per parlare di pendenza dobbiamo chiarire come ci muoviamo a partire dal punto, allora consideriamo una curva regolare che passa per il punto \underline{x}_0 :

$$\gamma : [-\epsilon, \epsilon] \subset \mathbb{R} \rightarrow A \subseteq \mathbb{R}^n, \quad \gamma(0) = \underline{x}_0$$

La curva rappresenta un sentiero nello spazio, definiamo pure il suo vettore velocità $\gamma'(0) = \underline{v}$, che indica la direzione iniziale del movimento. Visto che ci interessa sola la direzione (e non la velocità) scriviamo \underline{v} come il vettore di direzione unitaria, $\|\underline{v}\| = 1$.

Passo 2: Riduzione a una variabile

Definiamo la funzione composta (restrizione di f alla curva):

$$F(t) = (f \circ \gamma)(t) = f(\gamma(t))$$

Questa funzione descrive il valore percorrendo la curva γ . Poiché F dipende da una sola variabile, la sua derivata $F'(0)$ rappresenta la pendenza della funzione lungo il sentiero scelto nel punto $t = 0$ (corrispondente a \underline{x}_0).

Passo 3: La regola della catena

La differenziabilità di f ci garantisce che esiste una derivata unica in ogni direzione che approssima la funzione, ossia il gradiente. Allora possiamo applicare la regola della catena per calcolare

$$F'(0) = \langle \nabla f(\underline{x}_0), \gamma'(0) \rangle$$

Se poniamo il vettore unitario $\underline{v} = \gamma'(0)$, questa quantità coincide con la derivata direzionale di f in \underline{x}_0 lungo la direzione \underline{v}

$$\langle \nabla f(\underline{x}_0), \underline{v} \rangle = D_{\underline{v}} f(\underline{x}_0)$$

Dunque, la regola della catena ci dice che tutte le possibili pendenze di f sono governate dal prodotto scalare tra il gradiente e il vettore direzione.

Passo 4: Massimizzare la pendenza

Ricordando la definizione geometrica di prodotto scalare

$$\underline{a} \cdot \underline{b} = \|\underline{a}\| \|\underline{b}\| \cos \theta$$

possiamo scrivere:

$$D_{\underline{v}} f = \|\nabla f(\underline{x}_0)\| \cdot \underbrace{\|\underline{v}\|}_{=1} \cdot \cos \theta = \|\nabla f(\underline{x}_0)\| \cos \theta$$

dove θ è l'angolo tra il gradiente e il vettore unitario \underline{v} , che rappresenta la direzione scelta. Per massimizzare la pendenza $D_{\underline{v}} f$, dobbiamo massimizzare $\cos \theta$:

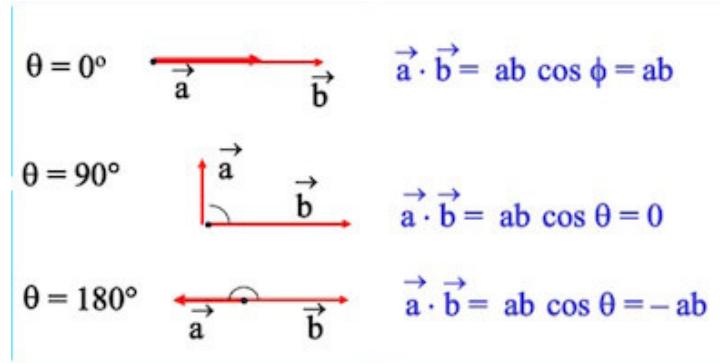
- **Massima Crescita ($\theta = 0^\circ$)**: La derivata è massima quando ci muoviamo nella direzione del gradiente,

$$D_{\underline{v}} f = \|\nabla f\|$$

- **Massima Decrescita ($\theta = \pi$)**: La derivata è minima (negativa) quando ci muoviamo in direzione opposta al gradiente,

$$D_{\underline{v}} f = -\|\nabla f\|$$

- **Crescita Nulla ($\theta = \pi/2$)**: La funzione non cambia valore se ci muoviamo ortogonalmente al gradiente (lungo le curve di livello).



4.3 Derivate successive

Sia $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ funzione derivabile su A , ossia esistono e sono ben definite le derivate

$$\frac{\partial f}{\partial x_1}(\underline{x}) \dots \frac{\partial f}{\partial x_n}(\underline{x})$$

Si suppone che $\forall f_{x_i}(\underline{x}) : A \rightarrow \mathbb{R}$ per qualche $i = 1 \dots n$ accade che la funzione $f_{x_i}(\underline{x})$ sia derivabile ulteriormente, allora esiste la derivata successiva di f in \underline{x}

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

La derivata seconda è, praticamente, la derivata parziale delle derivate prime di f e indicano la curvatura sulla superficie della funzione.

4.3.1 Matrice Hessiana e il teorema di Schwarz

Per raccogliere tutte le informazioni sulla curvatura della funzione, ci rifacciamo all'algebra lineare dove per ottenere uno scalare (un numero) moltiplicando dei vettori "due volte" (termini al quadrato tipo x^2 o misti xy), l'oggetto centrale deve essere necessariamente una matrice quadrata.

Nel caso generale, le derivate seconde di una funzione f sono date dal variare degli indici i e j tra 1 e n . Quindi saranno $n \times n$ derivate seconde possibili organizzate in una Matrice Hessiana:

$$H_f(\underline{x}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

In forma compatta si scrive anche

$$Hf(\underline{x}) = \begin{bmatrix} f_{x_1 x_1} & f_{x_1 x_2} & \cdots & f_{x_1 x_n} \\ f_{x_2 x_1} & f_{x_2 x_2} & \cdots & f_{x_2 x_n} \\ \vdots & \vdots & \ddots & \vdots \\ f_{x_n x_1} & f_{x_n x_2} & \cdots & f_{x_n x_n} \end{bmatrix}$$

Se $i = j$ la derivata si dice pura e si scrive

$$f_{x_i x_i}(\underline{x}) = \frac{\partial^2 f}{\partial x_i^2}(\underline{x})$$

Se, invece, $i \neq j$ la derivata si dice mista e si scrive

$$f_{x_i x_j}(\underline{x}) = \frac{\partial^2 f}{\partial x_i \partial x_j}(\underline{x})$$

Matrice hessiana nel caso di $n = 2$

Sia $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ e supponiamo che f sia derivabile 2 volte nell'insieme A , ossia $\forall (x, y) \in A$ esistono tutte e 4 le derivate seconde:

$$f_{xx}(x, y), f_{xy}(x, y), f_{yx}(x, y), f_{yy}(x, y)$$

Queste derivate seconde sono organizzate in una matrice hessiana di dimensione 2×2

$$Hf(x, y) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2}(x, y) & \frac{\partial^2 f}{\partial x \partial y}(x, y) \\ \frac{\partial^2 f}{\partial y \partial x}(x, y) & \frac{\partial^2 f}{\partial y^2}(x, y) \end{bmatrix}$$

E' fondamentale affermare che, nel caso generale, non possiamo dire che la matrice hessiana sia simmetrica, solo dopo le opportune analisi dimostrando che appartiene alla classe C^2 potremo verificarlo.

Il teorema di Schwarz

Teorema 4.3.1 (Teorema di Schwarz). *Sia $f : A \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$ una funzione definita su un aperto A . Sia $P_0 = (x_0, y_0) \in A$ e supponiamo che la funzione sia derivabile due volte su A , ossia che esistano $n \times n$ derivate seconde della forma $D_{x_i x_j}(\underline{x}) \quad \forall i, j = 1 \dots n \text{ e } \forall \underline{x} \in A$. Se le derivate miste f_{xy} e f_{yx} esistono in un intorno di P_0 e sono continue in P_0 , allora coincidono:*

$$f_{xy}(x_0, y_0) = f_{yx}(x_0, y_0)$$

Ovvero:

$$\frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) = \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0)$$

DIMOSTRAZIONE su $n = 2$

Sia $P = (x, y) \in A$ con $P \neq P_0$. Si definiscono le seguenti funzioni a partire da f :

$$F(x) = f(x, y) - f(x, y_0) \quad (\text{con } \forall y \text{ fissato})$$

$$G(y) = f(x, y) - f(x_0, y) \quad (\text{con } \forall x \text{ fissato})$$

$F(x)$ è la variazione della funzione sull'asse x quando y varia tra y e y_0 , mentre $G(y)$ è la variazione della funzione sull'asse y quando x varia tra x e x_0 .

Passo 1: Applicando il teorema del valor medio a $F(x)$ nell'intervallo di estremi x_0 e x , esiste un punto $x_1 \in (x_0, x)$ tale che:

$$F(x) - F(x_0) = F'(x_1)(x - x_0)$$

Poiché la derivata di F rispetto a x è data dalla differenza delle derivate parziali di f :

$$F'(x) = \frac{\partial f}{\partial x}(x, y) - \frac{\partial f}{\partial x}(x, y_0)$$

L'uguaglianza diventa:

$$F(x) - F(x_0) = \left(\frac{\partial f}{\partial x}(x_1, y) - \frac{\partial f}{\partial x}(x_1, y_0) \right) (x - x_0)$$

Applichiamo nuovamente il teorema del valor medio alla funzione $\frac{\partial f}{\partial x}$ (vista come funzione di y) nell'intervallo (y_0, y) . Esiste un punto $y_1 \in (y_0, y)$ tale che:

$$\frac{\partial f}{\partial x}(x_1, y) - \frac{\partial f}{\partial x}(x_1, y_0) = \frac{\partial^2 f}{\partial y \partial x}(x_1, y_1)(y - y_0)$$

Sostituendo, otteniamo l'espressione finale per la differenza di F :

$$F(x) - F(x_0) = \frac{\partial^2 f}{\partial y \partial x}(x_1, y_1)(y - y_0)(x - x_0)$$

Analogamente, si applica lo stesso ragionamento alla funzione $G(y)$ nell'intervallo di estremi y_0 e y . Esiste un punto $y_2 \in (y_0, y)$ tale che:

$$G(y) - G(y_0) = G'(y_2)(y - y_0) = \left(\frac{\partial f}{\partial y}(x, y_2) - \frac{\partial f}{\partial y}(x_0, y_2) \right) (y - y_0)$$

Applicando il teorema del valor medio rispetto a x alla funzione $\frac{\partial f}{\partial y}$ nell'intervallo (x_0, x) , esiste un punto $x_2 \in (x_0, x)$ tale che:

$$G(y) - G(y_0) = \frac{\partial^2 f}{\partial x \partial y}(x_2, y_2)(x - x_0)(y - y_0)$$

Passo 2: Osserviamo che le due differenze iniziali sono algebricamente identiche

$$F(x) - F(x_0) = [f(x, y) - f(x, y_0)] - [f(x_0, y) - f(x_0, y_0)]$$

$$G(y) - G(y_0) = [f(x, y) - f(x_0, y)] - [f(x, y_0) - f(x_0, y_0)]$$

Quindi $F(x) - F(x_0) = G(y) - G(y_0)$. Uguagliando le espressioni ottenute e dividendo per $(x - x_0)(y - y_0) \neq 0$:

$$\frac{\partial^2 f}{\partial y \partial x}(x_1, y_1) = \frac{\partial^2 f}{\partial x \partial y}(x_2, y_2)$$

Passo 3: Per l'ipotesi di continuità delle derivate miste in P_0 , passando al limite per $(x, y) \rightarrow (x_0, y_0)$, i punti intermedi (x_1, y_1) e (x_2, y_2) tendono entrambi a (x_0, y_0) :

$$\lim_{(x,y) \rightarrow (x_0,y_0)} \frac{\partial^2 f}{\partial y \partial x}(x_1, y_1) = \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0)$$

$$\lim_{(x,y) \rightarrow (x_0, y_0)} \frac{\partial^2 f}{\partial x \partial y}(x_2, y_2) = \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)$$

L'uguaglianza dei limiti dimostra che le derivate miste sono uguali e l'ordine di derivazione non importa:

$$\frac{\partial^2 f}{\partial y \partial x}(x_0, y_0) = \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)$$

Possiamo affermare che se $f \in C^2(A)$, la matrice Hessiana è simmetrica in ogni punto di A . □

La sola esistenza delle derivate miste in un punto non garantisce l'uguaglianza.

È fondamentale l'ipotesi di continuità delle derivate seconde.

Esistono controesempi (es. la funzione di Peano-Schwarz $f(x, y) = \frac{xy(x^2-y^2)}{x^2+y^2}$) dove le derivate miste nell'origine esistono ma sono diverse ($f_{xy} \neq f_{yx}$) proprio perché non sono continue in $(0, 0)$.

4.4 La formula di Taylor in funzioni multivariate

Per analizzare il comportamento di una funzione multivariata $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ in un punto, possiamo utilizzare la formula di Taylor per approssimare la funzione con un polinomio.

Per calcolare la formula di Taylor, è necessario ridurre il problema a una variabile, utilizzando una parametrizzazione che ci consenta di studiare il comportamento della funzione lungo un segmento.

Sia $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ che si suppone sia di classe $C^2(A)$ e siano $\underline{x}, \underline{x} + \underline{h} \in A \subseteq \mathbb{R}^n$ con $\underline{h} \neq 0$.

Possiamo dire che parametrizzare il segmento di estremi \underline{x} e $\underline{x} + \underline{h}$ contenuto in A :

$$[\underline{x}, \underline{x} + \underline{h}] = \{\underline{x}(t) \in \mathbb{R}^n : \underline{x}(t) = \underline{x} + t\underline{h}, t \in [0, 1]\} \subset A.$$

Si osserva che la parametrizzazione soddisfa le condizioni agli estremi:

$$\underline{x}(0) = \underline{x}, \quad \underline{x}(1) = \underline{x} + \underline{h}.$$

Ora possiamo scrivere la funzione ausiliaria (restrizione di f al segmento), sia $F : [0, 1] \rightarrow \mathbb{R}$:

$$F(t) = f(\underline{x} + t\underline{h}), \quad \forall t \in [0, 1].$$

Il nostro obiettivo, adesso, è quello di dimostrare che dalla formula di Taylor centrata in t possiamo ricavare informazioni su $f(\underline{x})$.

Se $f \in C^1(A)$, allora $F \in C^1([0, 1])$.

La derivata prima $F'(t)$ con la regola della catena:

$$F'(t) = \langle \nabla f(\underline{x} + t\underline{h}), \underline{h} \rangle = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\underline{x} + t\underline{h}) h_i.$$

Poiché abbiamo supposto che $f \in C^2(A)$, applicando nuovamente la regola della catena a ciascun termine di $\frac{\partial f}{\partial x_i}$ otteniamo:

$$F''(t) = \sum_{i,j=1}^n \frac{\partial f}{\partial x_j} \left(\frac{\partial f}{\partial x_i}(\underline{x} + t\underline{h}) \cdot h_i \right) \cdot h_j = \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(\underline{x} + t\underline{h}) \cdot h_i \cdot h_j$$

Formalmente, è un polinomio omogeneo di secondo grado con le componenti del vettore \underline{h} che appaiono moltiplicate fra di loro. I coefficienti di questo polinomio sono proprio le derivate seconde della funzione in $(\underline{x}, \underline{x} + t\underline{h})$:

- $F(0) = f(\underline{x}(0)) = f(\underline{x}) = \underline{x}$

- $F(1) = f(\underline{x}(1)) = f(\underline{x} + \underline{h}) = \underline{x} + \underline{h}$

Allora applicando la formula di Taylor alla $F(t)$ centrata nel punto $t = 0$ ottengo

$$F(t) = P_n(t, 0) + R_n(t, 0)$$

Generalizzando, se la funzione $f \in C^k(A)$ allora trasferisce la sua regolarità su una composta F . Possiamo affermare che $\exists \theta \in \mathbb{R}$ che sta tra $[0, 1]$ con resto tale che

$$F(1) = F(0) + F'(0) + \frac{F''(0)}{2!} + \cdots + \frac{F^{k-1}(0)}{k-1} + \underbrace{\frac{F^k(\theta)}{k!}}_{R_n}$$

Informalmente, invece di studiare il comportamento della funzione f nell'intero dominio multidimensionale $A \subseteq \mathbb{R}$, stiamo "affettando" il grafico della funzione lungo una retta. Definendo $F(t) = f(\underline{x} + t\underline{h})$, stiamo restringendo il nostro campo d'osservazione a ciò che accade esclusivamente lungo il segmento $(x, x + h)$.

I punti chiave di questa osservazione sono:

- **La Funzione come Binario**

La variabile scalare t agisce come un parametro temporale che ci fa scorrere lungo un binario rettilineo. Questo ci permette di ricondurre lo studio di una funzione di n variabili allo studio di una funzione di una sola variabile reale ($F : [0, 1] \rightarrow \mathbb{R}$);

- **La Derivata Direzionale**

Il risultato $F'(t) = \langle \nabla f(\underline{x} + t\underline{h}), \underline{h} \rangle$ non è altro che la derivata direzionale di f lungo il vettore \underline{h} (a meno di un fattore di normalizzazione). Ci dice quanto velocemente cambia il valore della funzione mentre ci muoviamo in quella specifica direzione \underline{h} , partendo dal punto $\underline{x} + t\underline{h}$.

- **Ponte verso i Teoremi di Taylor e del Valor Medio**

Questa costruzione è il ponte logico necessario per estendere i teoremi del calcolo monodimensionale a quello multidimensionale. Applicando il Teorema del Valor Medio o lo sviluppo di Taylor alla funzione ausiliaria $F(t)$ nell'intervallo $[0, 1]$, possiamo dimostrare rigorosamente le proprietà di f in \mathbb{R}^n , come l'esistenza di massimi/minimi o l'approssimazione quadratica tramite la matrice Hessiana.

4.4.1 La formula di Taylor secondo il resto di Lagrange e di Peano

Teorema 4.4.1 (Formula di Taylor con resto di Lagrange). *Sia $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ con A un aperto e $f \in C^k(A)$.*

Sia il segmento di estremi \underline{x} e $\underline{x} + \underline{h}$ contenuto in A , allora si può trovare un punto θ reale $\in [0, 1]$ tale che il polinomio di Taylor centrato in \underline{x} con resto secondo Lagrange di ordine 1 è:

$$f(\underline{x} + \underline{h}) = f(\underline{x}) + \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\underline{x} + \theta \underline{h}) h_i = f(\underline{x}) + \langle \nabla f(\underline{x} + \theta \underline{h}), \underline{h} \rangle$$

che equivale alla formula di Taylor $T_n \Rightarrow F(1) = F(0) + F'(0)$.

Nel caso in cui l'ordine sia l'ordine sia 2:

$$f(\underline{x} + \underline{h}) = f(\underline{x}) + \langle \nabla f(\underline{x} + \theta \underline{h}), \underline{h} \rangle + \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(\underline{x} + \theta \underline{h}) h_i h_j$$

che in forma compatta viene $f(\underline{x}) + \langle \nabla f(\underline{x}), \underline{h} \rangle + \frac{1}{2} \langle Hf(\underline{x} + \theta \underline{h}), \underline{h} \rangle$.

DIMOSTRAZIONE del caso $k = 2$

La matrice Hessiana in ogni punto di A è una matrice $n \times n$:

$$\langle Hf(\underline{x}) \cdot \underline{h} \rangle \text{ dove } \underline{h} \text{ è un vettore } (n \times 1)$$

In termini matriciali:

$$\underbrace{\begin{bmatrix} \dots \\ \dots \end{bmatrix}}_{n \times n} \cdot \underbrace{\begin{bmatrix} \vdots \\ \vdots \end{bmatrix}}_{n \times 1} \rightarrow \text{Vettore colonna}$$

Esplcitamente la matrice Hessiana è:

$$Hf(\underline{x} + \theta \underline{h}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(\underline{x} + \theta \underline{h}) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(\underline{x} + \theta \underline{h}) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(\underline{x} + \theta \underline{h}) & \dots & \frac{\partial^2 f}{\partial x_n^2}(\underline{x} + \theta \underline{h}) \end{bmatrix}$$

□

Diversa dalla formula con il resto di Lagrange è la formula con il resto di Peano perché si considera il gradiente e la matrice hessiana centrati in \underline{x} e non in un punto intermedio ignoto θ .

Teorema 4.4.2 (Formula di Taylor con resto di Peano). Sia $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ e $f \in C^2(A)$. Allora:

$$f(\underline{x} + \underline{h}) = f(\underline{x}) + \langle \nabla f(\underline{x}), \underline{h} \rangle + \frac{1}{2} \langle Hf(\underline{x}) \underline{h}, \underline{h} \rangle + o(\|\underline{h}\|^2) \quad \text{per } \underline{h} \rightarrow \underline{0}$$

DIMOSTRAZIONE del caso $n = 2$

Sia $f : A \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$. Consideriamo i 2 punti $(\underline{x}_0, \underline{y}_0)$ e l'incremento \underline{h} .

$$\underline{x} = (x_0, y_0), \quad \underline{h} = (h, k) \in \mathbb{R}^2$$

Quindi il punto incrementato è:

$$\underline{x} + \underline{h} = (x_0 + h, y_0 + k)$$

La formula si scrive:

$$f(x_0 + h, y_0 + k) = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)h + \frac{\partial f}{\partial y}(x_0, y_0)k + \frac{1}{2} \langle Hf(x_0, y_0) \underline{h}, \underline{h} \rangle$$

Analizziamo il termine quadratico (Hessiano). Poiché $f \in C^2(A)$, per il teorema di Schwarz la matrice è simmetrica ($f_{xy} = f_{yx}$):

$$Hf(x_0, y_0) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2}(x_0, y_0) & \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) \\ \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0) & \frac{\partial^2 f}{\partial y^2}(x_0, y_0) \end{bmatrix}$$

Questo va moltiplicato per $\underline{h} = \begin{pmatrix} h \\ k \end{pmatrix}$. Prima moltiplicazione (Matrice \times Vettore colonna):

$$\begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix} \begin{bmatrix} h \\ k \end{bmatrix} = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2}h + \frac{\partial^2 f}{\partial x \partial y}k \\ \frac{\partial^2 f}{\partial y \partial x}h + \frac{\partial^2 f}{\partial y^2}k \end{bmatrix}$$

Tutto questo va poi moltiplicato per \underline{h}^T (vettore riga) a sinistra:

$$\begin{bmatrix} h & k \end{bmatrix} \cdot \begin{bmatrix} f_{xx}h + f_{xy}k \\ f_{yx}h + f_{yy}k \end{bmatrix}$$

Svolgendo il prodotto:

$$\frac{1}{2} \left[\frac{\partial^2 f}{\partial x^2}(x_0, y_0)h^2 + \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)hk + \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0)hk + \frac{\partial^2 f}{\partial y^2}(x_0, y_0)k^2 \right]$$

Grazie alla simmetria ($f_{xy} = f_{yx}$), i termini centrali si sommano:

$$\text{Espressione omogenea di } 2^\circ \text{ grado in } h, k : \frac{\partial^2 f}{\partial x^2} h^2 + 2 \frac{\partial^2 f}{\partial x \partial y} hk + \frac{\partial^2 f}{\partial y^2} k^2$$

Tutto sommato per l'errore:

$$+o(\sqrt{h^2 + k^2}^2) = o(h^2 + k^2) \quad \text{con } h, k \rightarrow 0$$

Riscrivendo in modo opportuno ponendo $x = x_0 + h \implies h = x - x_0$ e $y = y_0 + k \implies k = y - y_0$:

$$P_2(x, y) = f(x_0, y_0) + \frac{\partial f}{\partial x}(x - x_0) + \frac{\partial f}{\partial y}(y - y_0) + \frac{1}{2} [f_{xx}(x - x_0)^2 + 2f_{xy}(x - x_0)(y - y_0) + f_{yy}(y - y_0)^2]$$

□

Questo è l'unico polinomio in x, y di grado ≤ 2 per cui si ha che la differenza tra $f(x, y)$ e P_2 va a zero più rapidamente della distanza al quadrato per i 2 punti $o\|((x - x_0)^2 + (y - y_0)^2)\|$.

$$\lim_{(x,y) \rightarrow (x_0,y_0)} \frac{f(x,y) - P_2(x,y)}{\|(x,y) - (x_0,y_0)\|^2} = 0$$

Nel caso n variabili (caso generale), il termine $\langle Hf(x_0)h, h \rangle$ "esce fuori" un'espressione omogenea di 2° grado in h_1, \dots, h_n , dove n è il numero di variabili, e si indicherà come $q(h)$, i cui coefficienti sono le opportune derivate seconde.

4.5 Studio di massimi e minimi per le funzioni a più variabili

Sia $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ una funzione definita su un aperto A e sia $\underline{x}_0 \in A$ un punto di A .

Definizione 4.5.1 (Massimo e minimo locale). *Se $\exists B(\underline{x}_0, r)$ con $r > 0$ (intorno di \underline{x}_0) per cui $\forall x \in A \cap B(\underline{x}_0, r)$ $f(\underline{x}_0) \leq f(x)$ allora \underline{x}_0 è un punto di minimo locale di f , mentre se $f(\underline{x}_0) \geq f(x)$ allora \underline{x}_0 è un punto di massimo locale di f .*

Si dice massimo (minimo) forte se la diseguaglianza è stretta, per $\underline{x} \neq \underline{x}_0$.

I punti di massimo e minimo relativi si chiamano anche punti di estremo locale.

4.6 Ottimizzazione libera

L'ottimizzazione libera si occupa di trovare i massimi e minimi di una funzione senza alcuna restrizione sul dominio. In questo caso, si cerca di identificare i punti in cui la funzione raggiunge i suoi valori estremi all'interno dell'intero spazio definito dalla funzione stessa.

Teorema 4.6.1 (Teorema di Fermat). *Sia f definita su un dominio $D = A \cup \partial A$ con A aperto, allora $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ e sia $\underline{x}_0 \in D$ o sia $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ e $\underline{x}_0 \in A$ un punto di estremo locale per f .*

Se f è differenziabile in tale punto \underline{x}_0 , il gradiente di f in \underline{x}_0 è nullo.

Questa è una condizione necessaria ma non sufficiente per l'esistenza di massimi e minimi locali perché esistono punti, detti critici o stazionari, che hanno gradiente nullo ma non sono né massimi né minimi (es. punti di sella o colle).

DIMOSTRAZIONE

Supponiamo che \underline{x}_0 sia un punto di massimo locale (la dimostrazione per il minimo è analoga).

Passo 1: Definizione della restrizione

Consideriamo la restrizione della funzione f lungo una direzione arbitraria governata dal versore \underline{v} .

Definiamo la funzione scalare di una sola variabile $F(t)$:

$$F(t) = f(\underline{x}_0 + t \cdot \underline{v})$$

dove \underline{v} è un versore e t varia in un intorno $(-\delta, \delta)$.

Passo 2: Osservazione sul massimo.

Poiché \underline{x}_0 è un punto di massimo locale per f , allora $f(\underline{x}_0) \geq f(\underline{x})$ per ogni \underline{x} in un intorno sferico.

Di conseguenza, in un intorno piccolo di t restringendoci alla retta:

$$F(0) = f(\underline{x}_0) \geq f(\underline{x}_0 + t \cdot \underline{v}) = F(t)$$

Quindi, $t = 0$ è un punto di massimo locale per la funzione $F(t)$.

Passo 3: Applicazione del Teorema di Fermat (Analisi I).

Applicando il classico Teorema di Fermat per funzioni di una variabile, sappiamo che la derivata nel punto di massimo interno deve annullarsi:

$$F'(0) = 0$$

Passo 4: Legame con le derivate parziali.

Calcoliamo $F'(t)$ usando la regola della catena:

$$F'(t) = \langle \nabla f(\underline{x}_0 + t \cdot \underline{v}), \underline{v} \rangle$$

Valutando in $t = 0$:

$$F'(0) = \langle \nabla f(\underline{x}_0), \underline{v} \rangle \iff \nabla f(\underline{x}_0) = \underline{0} \quad \forall \underline{v} \in \mathbb{R}^n$$

□

4.6.1 Test dell'hessiana

Per le funzioni ad una variabile, supposto che una funzione $f \in C^2$ e individuato un punto critico x_0 ($f'(x_0) = 0$), la classificazione del punto si ottiene dallo studio della derivata seconda, governata dal termine quadratico dello sviluppo di Taylor al secondo ordine.

Nel caso di funzioni di più variabili, l'analogo della derivata seconda è la matrice Hessiana.

Sia $f : A \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$ con A aperto e $f \in C^2(A)$, ossia esistono le 4 derivate seconde e sono continue.

Sia $P_0 = (x_0, y_0)$ un punto critico per f , ovvero $\nabla f(P_0) = \underline{0}$ e supponiamo che il determinante della matrice Hessiana nel punto non sia nullo: $\det Hf(x_0, y_0) \neq 0$.

$$\det Hf(x_0, y_0) = \det \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix} = f_{xx}f_{yy} - (f_{xy})^2$$

(Ricordando che per il Teorema di Schwarz $f_{xy} = f_{yx}$, quindi la matrice è simmetrica).

La classificazione avviene come segue:

- Se $\det H > 0$ e $f_{xx} > 0 \implies$ minimo locale;
- Se $\det H > 0$ e $f_{xx} < 0 \implies$ massimo locale;
- Se $\det H < 0 \implies$ punto di sella.

Se $\det H = 0$, il test è inefficace e bisogna procedere con altri metodi.

4.6.2 Forme Quadratiche

In un punto critico la parte lineare dello sviluppo di Taylor si annulla (il piano tangente è piatto). Per studiare il comportamento locale della funzione è quindi necessario analizzare il termine successivo, quello di secondo ordine. Quindi, per capire l'origine del test precedente, analizziamo la formula di Taylor al secondo ordine.

Il segno della variazione locale della funzione dipende dal termine di secondo grado:

$$f(\mathbf{x}_0 + \mathbf{h}) - f(\mathbf{x}_0) \approx \frac{1}{2} \underbrace{\langle Hf(\mathbf{x}_0) \mathbf{h}, \mathbf{h} \rangle}_{q(\mathbf{h})}$$

Il termine $q(\mathbf{h})$ è una forma quadratica.

Definizione 4.6.1 (Forma Quadratica). Una forma quadratica in $\mathbf{h} \in \mathbb{R}^n$ è un polinomio omogeneo di secondo grado nelle variabili h_1, \dots, h_n . Ad essa $q(\mathbf{h})$ è associata una matrice simmetrica A tale che:

$$q(\mathbf{h}) = \langle A\mathbf{h}, \mathbf{h} \rangle = \sum_{i,j=1}^n a_{ij} h_i h_j$$

Data la simmetria ($a_{ij} = a_{ji}$), i coefficienti del polinomio corrispondono agli elementi della matrice (i termini misti si dividono per 2).

Per esempio, dato il polinomio omogeneo:

$$q(\mathbf{h}) = 3h_1^2 + 10h_1h_2 + 27h_2^2$$

La matrice simmetrica associata A è:

$$A = \begin{bmatrix} 3 & 5 \\ 5 & 27 \end{bmatrix}$$

Notare come il coefficiente 10 di h_1h_2 sia stato diviso equamente tra a_{12} e a_{21} .

Per determinare la natura del punto critico, dobbiamo studiare il segno di $q(\mathbf{h})$ per $\mathbf{h} \neq \mathbf{0}$.

Sia A la matrice simmetrica reale, essa ammette n autovalori reali $\lambda_1, \dots, \lambda_n$.

- A è definita positiva: $q(\mathbf{h}) > 0, \forall \mathbf{h} \neq \mathbf{0}$.
 - $\lambda_i > 0$ per tutti gli i ;
 - Punto di minimo.
- A è definita negativa: $q(\mathbf{h}) < 0, \forall \mathbf{h} \neq \mathbf{0}$.
 - $\lambda_i < 0$ per tutti gli i ;
 - Punto di massimo.
- A è indefinita: $q(\mathbf{h})$ assume sia valori positivi sia negativi..
 - Esistono autovalori discordi (almeno un $\lambda > 0$ e un $\lambda < 0$).
 - Punto di sella.

Nel caso di 2 variabili, si può dire che gli autovalori λ_1, λ_2 sono le radici del polinomio caratteristico

$$\lambda^2 - \text{tr}(A)\lambda + \det(A) = 0$$

E valgono le relazioni

$$\det(A) = \lambda_1 \cdot \lambda_2 \quad ; \quad \text{tr}(A) = \lambda_1 + \lambda_2$$

Quindi la matrice simmetrica A di dimensione 2×2 associata alla forma quadratica $q(h, k)$ è:

$$A = \begin{bmatrix} a_{11} - \lambda & a_{12} \\ a_{12} & a_{22} - \lambda \end{bmatrix} = \lambda^2 - \underbrace{(a_{11} + a_{22})}_{\text{tr}(A)} \lambda + \underbrace{a_{11} \cdot a_{22} - a_{12}^2}_{\det(A)} = 0$$

Questo ci permette di spiegare il test dell'hessiana in modo più intuitivo:

1. Se $\det A > 0$, ovvero il prodotto $\lambda_1 \lambda_2$ è positivo \Rightarrow autovalori concordi.
 - Se $a_{11} > 0$ (o traccia > 0), allora sono entrambi positivi \rightarrow minimo.
 - Se $a_{11} < 0$ (o traccia < 0), allora sono entrambi negativi \rightarrow massimo.
2. Se $\det A < 0$, ovvero il prodotto $\lambda_1 \lambda_2$ è negativo \Rightarrow autovalori discordi \rightarrow sella.

Quindi, il problema geometrico della classificazione dei punti critici viene ricondotto a un problema algebrico: lo studio del segno della forma quadratica associata al termine di secondo ordine dello sviluppo di Taylor.

Possiamo riassumere il ragionamento in tre passaggi chiave:

1. Fermat non basta

Quando siamo in un punto critico ($\nabla f = 0$), il piano tangente è orizzontale. La parte lineare della funzione si annulla. Per capire se siamo su una cima, in una buca o su una sella, dobbiamo guardare il "termine successivo" dello sviluppo di Taylor: il termine di secondo grado (quadratico).

2. Gli autovalori sono le "curvature vere"

La matrice Hessiana descrive la curvatura, ma spesso contiene "termini misti" (f_{xy}) che creano confusione perché dipendono dal sistema di coordinate scelto (x, y) . Gli autovalori (λ) rappresentano le curvature lungo le direzioni principali, ovvero quelle direzioni speciali in cui la superficie curva senza torsioni.

- Un λ positivo indica una curvatura verso l'alto (convessità).
- Un λ negativo indica una curvatura verso il basso (concavità).

3. Hessiana vs Autovalori

A prima vista può sembrare strano: gli autovalori sono così importanti per classificare i punti critici, eppure quando scriviamo l'Hessiana

$$H = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{bmatrix}$$

non vediamo λ_1 e λ_2 da nessuna parte.

La chiave è capire che l'Hessiana "vede" la geometria con gli occhiali delle coordinate (x,y) scelte. Le derivate seconde f_{xx}, f_{xy}, f_{yy} descrivono come la funzione si curva lungo gli assi x e y e nelle direzioni miste. Ma queste direzioni sono arbitrarie: dipendono da come abbiamo orientato il nostro sistema di riferimento.

Gli autovalori, invece, sono invarianti geometrici: non dipendono dal sistema di coordinate. Essi rappresentano le curvature lungo le direzioni principali del paraboloida ed è qui che entra in gioco il teorema spettrale ci dice che esiste sempre una rotazione del sistema di coordinate (rappresentata da una matrice ortogonale Q) che allinea i nostri assi con le direzioni principali:

$$Q^T H Q = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = D$$

In questo nuovo sistema di coordinate (u,v) l'Hessiana diventa diagonale: i termini misti spariscono. La forma quadratica si semplifica a

$$q(u, v) = \lambda_1 u^2 + \lambda_2 v^2$$

Ora gli autovalori sono visibili: sono proprio i coefficienti sulla diagonale. In queste coordinate "speciali", la geometria diventa cristallina:

- $\lambda_1 > 0, \lambda_2 > 0$ la parabola che sale in entrambe le direzioni \Rightarrow minimo;
- $\lambda_1 < 0, \lambda_2 < 0$ la parabola che scende in entrambe le direzioni \Rightarrow massimo;
- $\lambda_1 \cdot \lambda_2 < 0$ la parabola sale in una direzione e scende nell'altra \Rightarrow sella

4.6.3 Caso dubbio: il test dell'hessiana non funziona

Quando il determinante è nullo, la matrice Hessiana è "degenera" e non ci dà informazioni sufficienti sulla curvatura (la funzione è "troppo piatta" per l'approssimazione quadratica).

Infatti, il metodo dell'Hessiana è inconcludente se abbiamo almeno un autovalore uguale a zero: non possiamo capire in modo certo se quel punto è di massimo, di minimo o di sella, e dobbiamo quindi ricorrere a un metodo alternativo.

Nell'analisi unidimensionale possiamo immaginare di voler classificare un punto critico x_1 di una funzione $f(x)$ sapendo solo che $f(x_1) = c$. Guardando il grafico sapremmo riconoscere immediatamente se x_1 è un massimo o un minimo.

Senza accesso al grafico, possiamo traslare verticalmente la funzione in modo che l'estremo si trovi esattamente sull'asse delle x , definiamo quindi

$$\bar{f}(x) = f(x) - c$$

Questa nuova funzione è identica a quella originale, ma "abbassata" (o "innalzata", se $c < 0$) esattamente del valore c . In particolare:

$$\bar{f}(x_1) = f(x_1) - c = 0$$

Ora studiamo il segno di \bar{f} in un intorno di x_1 :

- Se $\bar{f}(x) \leq 0$ in un intorno di x_1 , significa che dopo la traslazione tutti i valori della funzione vicino a x_1 sono negativi o nulli.
Quindi $f(x) \leq c$ in quell'intorno $\implies x_1$ è un punto di massimo.
- Se $\bar{f}(x) \geq 0$ in un intorno di x_1 , significa che $f(x) \geq c$ in quell'intorno $\implies x_1$ è un punto di minimo.
- Se \bar{f} cambia segno, allora x_1 non è né un massimo né un minimo (caso che in una variabile è raro, ma in più variabili corrisponde alla sella).

Questo stesso ragionamento si applica alle funzioni di due variabili quando l'Hessiana è degenera. In questo caso, si deve tornare alla definizione di estremo locale, studiando direttamente il segno della variazione della funzione in un intorno del punto critico $P_0 = (x_0, y_0)$

Si definisce l'incremento:

$$\Delta f(x, y) = f(x, y) - f(x_0, y_0)$$

Questa è esattamente l'analogo bidimensionale della funzione \bar{f} vista prima: stiamo "traslando" la superficie in modo che il punto critico si trovi a quota zero, e poi studiamo se la superficie nei dintorni è tutta sopra, tutta sotto, o a cavallo del piano $z = 0$.

Si studia quindi il segno di Δf in un intorno sufficientemente piccolo di P_0 :

1. **Minimo Locale:** Se esiste un intorno in cui $\Delta f(x, y) \geq 0$ per ogni punto (ovvero la funzione assume sempre valori maggiori o uguali a $f(P_0)$);
2. **Massimo Locale:** Se esiste un intorno in cui $\Delta f(x, y) \leq 0$ per ogni punto (la superficie "traslata" è tutta sotto il piano $z = 0$);
3. **Punto di Sella:** Se in qualsiasi intorno di P_0 , per quanto piccolo, esistono sia punti in cui $\Delta f > 0$ sia punti in cui $\Delta f < 0$. Graficamente, studiando il segno sul piano xy , vedremo delle regioni positive e negative che si alternano e convergono tutte nel punto critico.

4.7 Ottimizzazione vincolata

A differenza dell'ottimizzazione libera, dove cerchiamo estremi su un aperto $A \subseteq \mathbb{R}^n$, qui cerchiamo i massimi e minimi di una funzione $f(x, y)$ limitatamente ai punti che appartengono a un insieme V (detto *vincolo*).

$$V = \{(x, y) \in \mathbb{R}^2 : \underbrace{g(x, y) = k}_{g(x, y) - k = 0}\} \subseteq \mathbb{R}^2 \quad \text{con } g \in C^1$$

In genere V è una curva di livello (es. una circonferenza, una retta) definita implicitamente da un'equazione $g(x, y) = 0$ (o k).

Quindi risulta possibile esplicitare una funzione rispetto ad un'altra (teorema della funzione implicita):

$y = y(x)$ allora il vincolo diventa $g(x, y(x))$, una funzione ad una variabile

Se si studia la natura dei punti critici quando l'insieme di definizione è compatto, si può usare il Teorema di Weierstrass.

Teorema 4.7.1 (Teorema di Weierstrass). *Sia $f : K \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ una funzione continua definita su un insieme compatto $K \subset \mathbb{R}^n$, ossia K è un sottoinsieme chiuso e limitato di \mathbb{R}^n . Allora f assume sia il massimo che il minimo assoluto su in almeno due punti di K .*

Si può anche definire il teorema come:

$$\exists (x_{\min}, y_{\min}), (x_{\max}, y_{\max}) \in K : f(x_{\min}, y_{\min}) \leq f(x, y) \leq f(x_{\max}, y_{\max}) \quad \forall (x, y) \in K$$

4.7.1 Parametrizzazione

Se l'equazione del vincolo $g(x, y) = 0$ è semplice, possiamo parametrizzare la curva $(x(t), y(t))$ trasformando il vincolo in un sostegno ad una curva $V = (\gamma[a, b])$.

$$\begin{cases} x = x(t) \\ y = y(t) \end{cases} \quad t \in [0, h]$$

Questo metodo è comodo per vincoli lineari (es. $3x + 5y = 7$) ma difficile per vincoli complessi.

4.7.2 Moltiplicatori di Lagrange

Siano due funzioni $f, g : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ e siano entrambe appartenenti alla classe $C^1(A)$. e sia l'insieme $V = \{(x, y) \in A : g(x, y) = k\}$.

Supponendo che $\underline{x_0} \in V$ sia un punto di estremo locale per f ristretto all'insieme V e che il gradiente $\nabla g(\underline{x_0}) \neq \underline{0}$ (regolare rispetto in quel punto), allora esiste un numero reale $\lambda \in \mathbb{R}$ (detto moltiplicatore di lagrange) tale che:

$$\nabla f(x, y) = \lambda \nabla g(x, y)$$

Geometricamente, significa che nel punto di estremo le curve di livello di f sono tangenti al vincolo V (i gradienti sono paralleli).

Per trovare i candidati punti di estremo vincolato definiamo una funzione detta Lagrangiana:

$$\mathcal{L}(x, y, \lambda) = f(x, y) - \lambda[g(x, y) - k]$$

che ha la stessa regolarità della funzione di cui è composta.

I punti critici vincolati si trovano cercando i punti stazionari di \mathcal{L} (dove il gradiente "completo" si annulla). Si risolve il sistema:

$$\nabla \mathcal{L}(x, y, \lambda) = \underline{0} \Leftrightarrow \begin{cases} \frac{\partial f}{\partial x} - \lambda g_x = 0 \\ \frac{\partial f}{\partial y} - \lambda g_y = 0 \\ g(x, y) - k = 0 \quad (\text{la derivata del vincolo rispetto a } \lambda) \end{cases}$$

Le soluzioni (x, y, λ) ci forniscono i candidati (x, y) da valutare, che vanno poi testati con altri metodi per trovare i massimi e minimi assoluti.

4.7.3 Esempio di ottimizzazione vincolata con Sostituzione e Lagrange.

Trovare max e min di $f(x, y) = (x - 1)^2 - y^2$ soggetta al vincolo $x^2 + y^2 = 1$.

Il vincolo è la circonferenza unitaria (compatto \rightarrow Weierstrass garantisce l'esistenza).

Metodo 1 (Sostituzione): Dal vincolo: $y^2 = 1 - x^2$ (con $x \in [-1, 1]$). Sostituiamo in f :

$$h(x) = (x - 1)^2 - (1 - x^2) = x^2 - 2x + 1 - 1 + x^2 = 2x^2 - 2x$$

Derivata: $h'(x) = 4x - 2 \implies x = 1/2$. Confrontiamo i valori in $x = 1/2$ e agli estremi $x = \pm 1$:

- $x = 1 \implies y = 0 \implies f(1, 0) = 0$
- $x = -1 \implies y = 0 \implies f(-1, 0) = 4$ (Max)
- $x = 1/2 \implies y = \pm\sqrt{3}/2 \implies f(1/2, \pm\sqrt{3}/2) = -1/2$ (Min)

Metodo 2 (Lagrange): Costruiamo $\mathcal{L}(x, y, \lambda) = (x - 1)^2 - y^2 - \lambda(x^2 + y^2 - 1)$. Sistema:

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

Dalla seconda equazione: $y = 0$ o $\lambda = -1$.

- Se $y = 0$: dal vincolo $x = \pm 1$. Ottengo i punti $(1, 0)$ e $(-1, 0)$.
- Se $\lambda = -1$: dalla prima equazione $2(x - 1) = -2x \implies 2x - 2 = -2x \implies 4x = 2 \implies x = 1/2$. Dal vincolo $y^2 = 1 - (1/4) = 3/4 \implies y = \pm\sqrt{3}/2$.

Si ritrovano gli stessi 4 punti candidati.

4.8 Integrazione

Mentre l'integrale dell'analisi unidimensionale $\int_a^b f(x)dx$ calcola l'area tra la curva e l'asse x , l'integrale doppio $\iint_R f(x, y) dx dy$ calcola il volume compreso tra la superficie definita da $z = f(x, y)$ e il piano xy sopra un dominio R nel piano.

Sia una funzione $f : R \rightarrow \mathbb{R}$ definita su un rettangolo $R = [a, b] \times [c, d]$, supponiamo che D_1 e D_2 siano due partizioni di $[a, b]$ e $[c, d]$ rispettivamente definite da $n + 1$ e $m + 1$ punti:

$$D_1 = \{a = x_0, x_1, \dots, x_n = b\} \quad D_2 = \{c = y_0, y_1, \dots, y_m = d\}$$

Consideriamo $D = D_1 \times D_2$ un dominio nel piano xy , esso ci dà una suddivisione del rettangolo in r_{ij} , dove si definisce $r_{ij} = [x_{i-1}, x_i] \times [y_{j-1}, y_j]$.

Il rettangolo R viene quindi suddiviso in $n \cdot m$ sotto-rettangoli r_{ij} .

Ora, supponiamo che f sia limitata sul rettangolo R , tale che esistono due numeri reali m e M per cui $m \leq f(x, y) \leq M$ per ogni $(x, y) \in R$.

Se per ogni rettangolino considero $m_{i,j} = \inf_{(x,y) \in Q_{ij}} f(x, y)$ come l'estremo inferiore (il valore minimo) assunto dalla $f(x, y)$, il volume del parallelepipedo sotto la superficie è

$$\sum_{i,j=1}^{n,m} m_{i,j} \cdot \underbrace{(x_i - x_{i-1}) \cdot (y_j - y_{j-1})}_{\text{Area}(R_{i,j})}$$

La somma di tutti questi volumi viene definita come la somma inferiore: $s = (f, D)$.

Stesso ragionamento vale per la somma superiore $S(f, D)$, dove si considera $M_{i,j} = \sup_{(x,y) \in Q_{ij}} f(x, y)$ come l'estremo superiore (il valore massimo) assunto dalla $f(x, y)$.

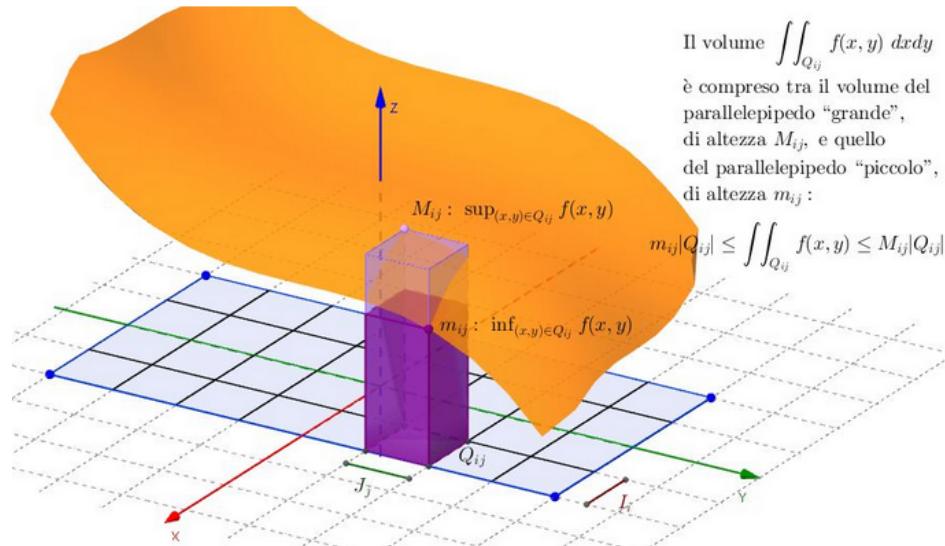
$$\sum_{i,j=1}^{n,m} M_{i,j} \cdot \underbrace{(x_i - x_{i-1}) \cdot (y_j - y_{j-1})}_{\text{Area}(R_{i,j})}$$

Queste due somme sono legate dalla relazione proveniente dal teorema della media integrale

$$m(b-a)(d-c) \leq s \leq S \leq M(b-a)(d-c) \quad \forall \text{ suddivisione di } R$$

Al variare di D , le somme inferiori e superiori variano, ma rimangono sempre limitate e ben definite tra $m(b-a)(d-c)$ e $M(b-a)(d-c)$.

Quindi, come nel caso unidimensionale si dimostra che vale la relazione per cui $\sup_D s(f, D) \leq \inf_D S(f, D)$, che nel caso di una uguaglianza ci permette di definire una funzione integrabile.



4.8.1 Integrale doppio e concetto di continuità

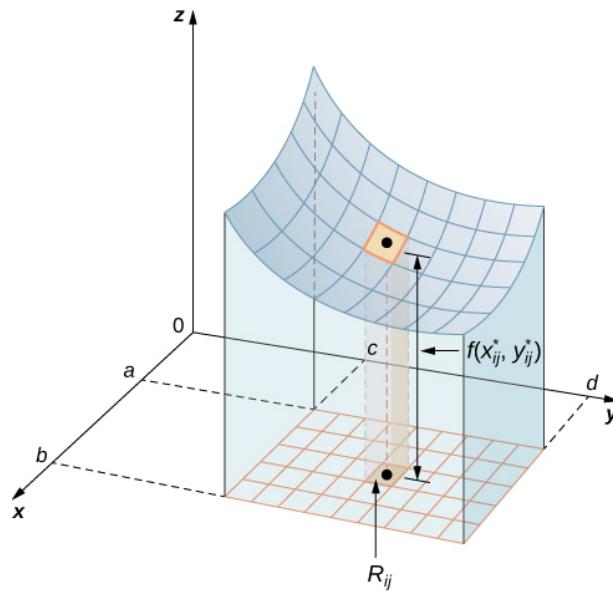
Definizione 4.8.1 (Integrale Doppio). Una funzione $f : R \rightarrow \mathbb{R}$ limitata su R dove $R = [a, b] \times [c, d]$. La funzione f si dice integrabile secondo Riemann su R se

$$\sup_D s(f, D) = \inf_D S(f, D)$$

In tal caso, il valore comune(un numero reale) viene detto integrale doppio di f su R e si indica con

$$\iint_R f, \quad I(f, R), \quad \iint_R f(x, y) dx dy$$

Se si ha che la funzione $f(x, y) \geq 0$ per ogni $(x, y) \in R$, l'integrale doppio rappresenta il volume della regione solida compresa tra la superficie $z = f(x, y)$ e il piano xy sopra il rettangolo R , ossia un cilindroide curvilineo.



Ogni addendo nelle somme inferiori e superiori è il volume di un parallelepipedo di base il rettangolo r_{ij} e altezza rispettivamente $m_{i,j}$ e $M_{i,j}$.

Teorema 4.8.1 (Teorema sulla continuità). Sia R un rettangolo chiuso e limitato in \mathbb{R}^2 definito come $R = [a, b] \times [c, d]$ e sia $f : R \rightarrow \mathbb{R}$ una funzione a più variabili. Se la funzione è continua su tutto il rettangolo R , allora f è integrabile secondo Riemann su R .

4.8.2 Calcolo Integrale sui rettangoli

Sia $f : R \rightarrow \mathbb{R}$ una funzione continua definita su un rettangolo $R = [a, b] \times [c, d]$, se $f(x, y) = g(x) \cdot h(y)$ con $g : [a, b] \rightarrow \mathbb{R}$ e $h : [c, d] \rightarrow \mathbb{R}$ integrabili su i rispettivi intervalli, allora f è integrabile su R e vale la seguente proprietà di separazione:

$$\iint_R f(x, y) dx dy = \left(\int_a^b g(x) dx \right) \cdot \left(\int_c^d h(y) dy \right)$$

Questo è il caso semplice in cui la funzione si può "separare" in due funzioni di una sola variabile.

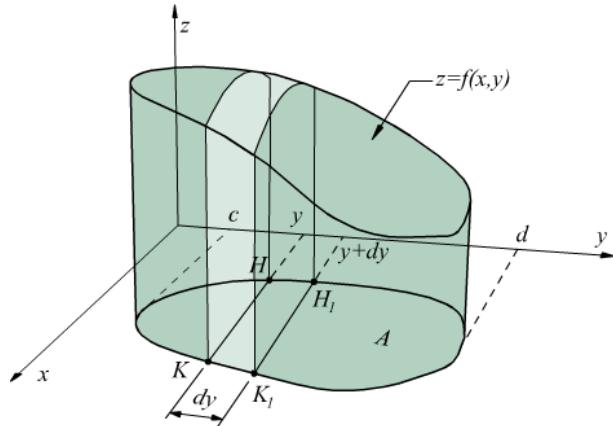
Teorema di Fubini-Tonelli

Nel caso generale, si definisce il teorema di Fubini-Tonelli, che permette di calcolare l'integrale doppio come integrali singoli successivi (integrali iterati). Sia $f : R \rightarrow \mathbb{R}$ una funzione continua

definita su un rettangolo $R = [a, b] \times [c, d]$, allora f è integrabile su R e vale la seguente proprietà di riduzione:

$$\iint_R f(x, y) dx dy = \int_a^b \left[\int_c^d f(x, y) dy \right] dx = \int_c^d \left[\int_a^b f(x, y) dx \right] dy$$

Quindi, prima integro rispetto ad una variabile (tenendo l'altra fissa), e poi integro il risultato rispetto all'altra variabile. Geometricamente significa "affettare" il volume lungo un asse e poi sommare le aree delle fette. L'ordine di integrazione è indifferente (per funzioni continue).



4.8.3 Integrali doppi su domini generali e la misurabilità secondo Peano-Jordan

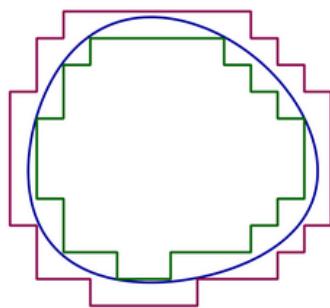
Sia D una regione del piano limitata e sia f definita e limitata su questo dominio. Estendiamo la funzione f a tutto il rettangolo R che contiene D , ponendo $f(x, y) = 0$ per ogni punto $(x, y) \notin D$.

Definizione 4.8.2. Sia f una funzione definita e limitata su un dominio limitato $D \subset \mathbb{R}^2$. La funzione f si dice integrabile secondo Riemann su D se la funzione estesa \tilde{f} è integrabile secondo Riemann sul rettangolo R che contiene D . In tal caso, l'integrale doppio di f su D è definito come:

$$\iint_D f(x, y) dx dy = \iint_R \tilde{f}(x, y) dx dy$$

Non tutti gli insiemi possono essere misurabili perché non siamo in grado di dire con certezza quanto è grande la sua area. E se non conosciamo l'area di base non possiamo calcolare il volume che si sta sopra. Un sottoinsieme limitato $D \subset \mathbb{R}^2$ è misurabile secondo Peano-Jordan se la funzione f (identicamente 1 su D e 0 fuori) è integrabile secondo Riemann sul rettangolo che contiene D . In tale caso

$$|D| = \iint_D 1 dx dy \implies \text{area di } D$$



L'idea di misurabilità secondo Peano-Jordan: se riesco a "riempire" l'insieme con dei rettangoli (o a "svuotarlo" con dei rettangoli) in modo che le somme inferiori e superiori coincidano, allora posso definire la sua area.

In generale, ogni rettangolo $R = [a, b] \times [c, d]$ è misurabile secondo Peano-Jordan, e la sua misura è data da:

$$|R| = (b - a)(d - c) \text{ che è l'area del rettangolo.}$$

Tutti gli insiemi che conosciamo dalla geometria elementare (poligoni, cerchi, ecc.) sono misurabili secondo Peano-Jordan e la loro misura corrisponde alla formula delle loro aree.

Un esempio di un poligono che non è misurabile secondo Peano-Jordan

Sia Q un quadrato $[1, 0] \times [0, 1]$ con coordinate razionali e sia una funzione $f : Q \rightarrow \mathbb{R}$ definita come:

$$f(x, y) = \begin{cases} 1 & \text{se } x, y \in \mathbb{Q} \\ 0 & \text{altrimenti} \end{cases} \quad \text{Funzione di Dirichlet}$$

Questa funzione non è integrabile secondo Riemann su Q , perché in ogni rettangolino r_{ij} , per quanto piccolo, la funzione assume sia il valore 0 che il valore 1 (i numeri razionali sono densi nei reali). Quindi non esiste la somma inferiore uguale alla somma superiore, e di conseguenza l'insieme non è misurabile secondo Peano-Jordan.

4.8.4 Proprietà degl integrali doppi

Analogamente agli integrali singoli:

- **Linearità:** L'integrale di una combinazione lineare è la combinazione lineare degli integrali.

$$\iint_D (\alpha f + \beta g) = \alpha \iint_D f + \beta \iint_D g$$

- **Additività sul dominio:** Se D viene diviso in due sottodomini D_1 e D_2 che non si sovrappongono (se non sul bordo), l'integrale su D è la somma degli integrali.

- **Monotonia:** Se $f(x, y) \leq g(x, y)$ su tutto D , allora $\iint_D f \leq \iint_D g$.

- **Relazione della funzione con il suo valore assoluto:** Se f è integrabile su D , allora $|f|$ è integrabile su D e vale:

$$\left| \iint_D f(x, y) dx dy \right| \leq \iint_D |f(x, y)| dx dy$$

E se D è misurabile secondo Peano-Jordan, allora $|D| = \iint_D 1 dx dy$ e posso così definire:

$$\left| \iint_D f(x, y) dx dy \right| \leq \iint_D |f(x, y)| dx dy \leq \underbrace{\sup_{(x,y) \in D} |f(x, y)|}_{\text{massimo valore assoluto di } f \text{ in } D} \cdot |D|$$

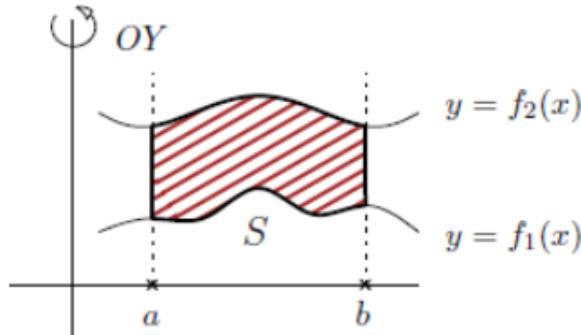
4.8.5 Calcolo Integrale su domini semplici

Non tutti i domini di integrazione sono rettangoli. Spesso ci troviamo a dover integrare su figure dove gli estremi dipendono dalla posizione. Si dicono domini normali (o semplici) quelle regioni del piano delimitate da due rette parallele e dai grafici di due funzioni continue. Ne esistono di due tipi:

- Dominio normale rispetto all'asse x (**Y-simplice**):

È un insieme $D \subset \mathbb{R}^2$ che si proietta sull'asse x in un intervallo chiuso $[a, b]$. Le sue "pareti" superiore e inferiore sono date da due funzioni continue $f_1(x)$ e $f_2(x)$.

$$D = \{(x, y) \in \mathbb{R}^2 : a \leq x \leq b, f_1(x) \leq y \leq f_2(x)\}$$



D è l'unione di tutti i segmenti verticali compresi tra le curve $y = f_1(x)$ e $y = f_2(x)$ per ogni x in $[a, b]$ e si può scrivere come:

$$D = \bigcup_{x \in [a, b]} \{x\} \times [f_1(x), f_2(x)]$$

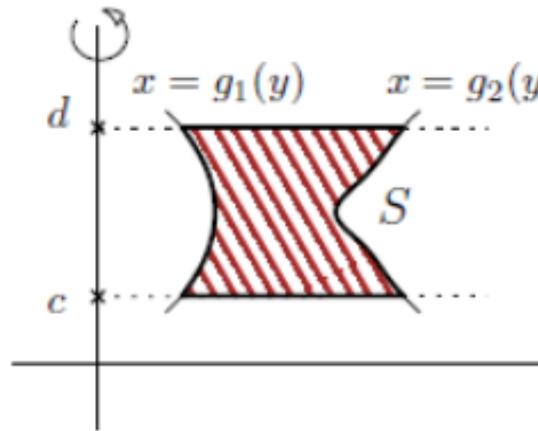
- Dominio normale rispetto all'asse y (**X-simplice**):

È un insieme D che si proietta sull'asse y in un intervallo $[c, d]$. Le sue pareti destra e sinistra sono funzioni della y , diciamo $h_1(y)$ e $h_2(y)$.

$$D = \{(x, y) \in \mathbb{R}^2 : c \leq y \leq d, h_1(y) \leq x \leq h_2(y)\}$$

D è l'unione di tutti i segmenti orizzontali compresi tra le curve $x = h_1(y)$ e $x = h_2(y)$ per ogni y in $[c, d]$ e si può scrivere come:

$$D = \bigcup_{y \in [c, d]} [h_1(y), h_2(y)] \times \{y\}$$



Le misure dei due insiemmi, pensando all'interpretazione geometrica, sono l'area delle regioni che si ottengono:

$$D_{y-\text{simplice}} = \int_a^b [f_2(x) - f_1(x)] dx$$

$$D_{x-\text{semplice}} = \int_c^d [h_2(y) - h_1(y)] dy$$

Teorema 4.8.2 (Formule di riduzione per domini normali). *Ogni funzione continua $f : D \rightarrow \mathbb{R}$ definita su un dominio normale D è integrabile secondo Riemann su D e vale:*

- *Se D è Y -semplice:*

$$\iint_D f(x, y) dx dy = \int_a^b \left(\int_{f_1(x)}^{f_2(x)} f(x, y) dy \right) dx$$

- *Se D è X -semplice:*

$$\iint_D f(x, y) dx dy = \int_c^d \left(\int_{h_1(y)}^{h_2(y)} f(x, y) dx \right) dy$$

Supponiamo di avere D_1, D_2, \dots, D_k domini semplici sul piano e supponiamo che a due a due non hanno punti in comune tranne che sui bordi. Se f è una funzione continua sull'unione $D = \bigcup_{i=1}^k D_i$, allora f è integrabile su D e vale la seguente proprietà di additività:

$$\iint_D f(x, y) dx dy = \sum_{i=1}^k \iint_{D_i} f(x, y) dx dy$$

D è una regione composta da più domini semplici, decomponibile.

4.8.6 Domini simmetrici e la jacobiana

Nell'analisi unidimensionale il cambio di variabile è possibile se la funzione è di classe C^1 e biunivoca (invertibile):

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

Nell'analisi multidimensionale, spesso quando abbiamo domini che presentano delle simmetrie ci conviene cambiare le variabili trasformando il dominio D (nel piano xy) in un dominio D' (nel piano ρ, θ) molto più semplice (ad esempio un rettangolo), tramite una trasformazione biunivoca. Sia $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ tale che

$$(\rho, \theta) \rightarrow f(\rho, \theta) = (f_1(\rho, \theta), f_2(\rho, \theta)) = (x(\rho, \theta), y(\rho, \theta))$$

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

La trasformazione deve essere biunivoca e di classe C^1 (derivabile con continuità). Questo ci permette di definire la matrice Jacobiana della trasformazione:

$$J_f(\rho, \theta) = \begin{bmatrix} \nabla f_1(\rho, \theta) \\ \nabla f_2(\rho, \theta) \end{bmatrix} = \begin{bmatrix} \frac{\partial x}{\partial \rho} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial \rho} & \frac{\partial y}{\partial \theta} \end{bmatrix} = \begin{bmatrix} \cos \theta & -\rho \sin \theta \\ \sin \theta & \rho \cos \theta \end{bmatrix}$$

Il determinante Jacobiano è quindi:

$$\det(J_f(\rho, \theta)) = \rho \cos^2 \theta - (-\rho \sin^2 \theta) = \rho(\cos^2 \theta + \sin^2 \theta) = \rho$$

(per la proprietà trigonometrica fondamentale $\cos^2 \theta + \sin^2 \theta = 1$).

Se il determinante è diverso da zero, la trasformazione è localmente invertibile.

Con $\rho = 0 \implies x = 0, y = 0$ (l'origine), la trasformazione non è biunivoca. Inoltre, data la periodicità di 2π , per renderla biunivoca restringiamo il dominio di θ a $[0, 2\pi)$ (o $(-\pi, \pi]$) e $\rho > 0$.

Consideriamo quindi l'aperto $A = (0, +\infty) \times (0, 2\pi)$.

Teorema 4.8.3 (Cambiamento di variabili). Siano D e T due insiemi aperti misurabili del piano. Sia $F : T \rightarrow D$ una trasformazione di coordinate (cambiamento di variabili) data da $F(\rho, \theta) = (x(\rho, \theta), y(\rho, \theta))$.

Supponiamo che:

- F sia biunivoca e di classe C^1 sull'interno di T ;
- Il determinante Jacobiano $\det J_{F(\rho, \theta)}$ sia diverso da zero ovunque in T .

Allora, per ogni funzione $f(x, y)$ integrabile su D , vale la formula:

$$\iint_D f(x, y) dx dy = \iint_T f(\rho \cos \theta, \rho \sin \theta) \cdot \rho d\rho d\theta$$

