

**Corso di Analisi 2**

**Domande e risposte al**

**macro-argomento:**

**CALCOLO DIFFERENZIALE**

Ultima modifica: 31 gennaio 2026

# Indice

<b>1 Calcolo differenziale</b>	<b>3</b>
1.1 Funzione a più variabili e campo di esistenza . . . . .	3
1.1.1 Grafico di funzione e sezioni (trasversali e orizzontali) . . . . .	3
1.1.2 Curve . . . . .	4
1.2 Limiti e continuità . . . . .	5
1.2.1 Continuità . . . . .	7
1.2.2 Coordinate Polari . . . . .	8
1.3 Derivata . . . . .	11
1.3.1 Derivata parziale . . . . .	11
1.3.2 Gradiente . . . . .	12
1.4 Differenziabilità . . . . .	14
1.5 Derivate successive . . . . .	23
1.5.1 Matrice Hessiana . . . . .	24
1.6 La formula di Taylor in funzioni multivariate . . . . .	26
1.7 Studio dei massimi e minimi di funzioni multivariate . . . . .	29
1.8 Ottimizzazione libera . . . . .	30
1.8.1 Test dell'hessiana . . . . .	30
1.8.2 Forme Quadratiche . . . . .	31
1.8.3 Caso dubbio: il test dell'hessiana non funziona . . . . .	33
1.9 Ottimizzazione vincolata . . . . .	34
1.9.1 Parametrizzazione . . . . .	34
1.9.2 Moltiplicatori di Lagrange . . . . .	34
1.10 Integrazione . . . . .	36
1.10.1 Calcolo Integrale sui rettangoli . . . . .	37
1.10.2 Integrali doppi su domini generali . . . . .	37
1.10.3 Calcolo Integrale sulle regioni semplici . . . . .	38
1.10.4 Cambiamento di variabili con integrali doppi . . . . .	40

# Calcolo differenziale

## 1.1 Funzione a più variabili e campo di esistenza

### 59. Cosa si intende per una funzione a più variabili?

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.

### 60. Cos'è il dominio di una funzione a più variabili?

**Definizione 1.1.1 (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  $\mathbb{R}^2$ .

### 61. Cosa si intende l'immagine di una funzione a più variabili?

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

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

### 1.1.1 Grafico di funzione e sezioni (trasversali e orizzontali)

### 62. Cosa si intende per grafico di una funzione a più variabili?

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

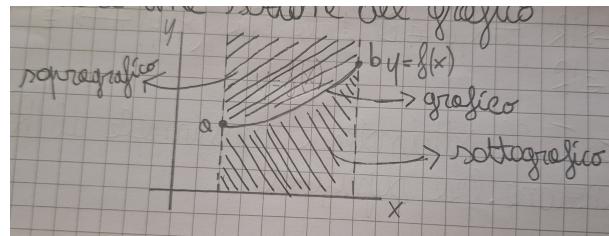
Geometricamente,  $\forall (x, y) \in D$  consideriamo un punto  $(x, y, f(x, y))$  nello spazio, dove  $f(x, y)$  è il valore della funzione.

### 63. Come si definisce un insieme del piano "scritto" come un'equazione o una disequazione?

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.

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



#### 64. Cosa sono le sezioni di un grafico a piani e le sezioni trasversali?

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 1.1.3 (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.

#### 65. Cosa sono gli insiemi di livello (o curve di livello) di una funzione di due variabili?

**Definizione 1.1.4 (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.

### 1.1.2 Curve

#### 66. Cos'è una curva?

**Definizione 1.1.5 (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  $\gamma_i$  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  $\mathbb{R}^n$

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

### 67. Cosa sono le curve piane, cartesiane o chiuse?

**Definizione 1.1.6 (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 1.1.7 (Curva Cartesiana).** E' una curva che nasce direttamente dal grafico della funzione scalare  $z = f(x)$ . Ogni funzione  $f : [a, b] \rightarrow \mathbb{R}$  può essere vista come una curva parametrizzando  $x$  con  $t$ .

$$\gamma(t) = (t, f(t))$$

con  $t \in [a, b]$

Le curve cartesiane sono sempre regolari.

**Definizione 1.1.8 (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  $\gamma(a) = \gamma(b)$ .

### 68. Come si può riformulare per le funzioni multivariate il teorema degli zeri?

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  $A_i \subseteq A$  tale che  $f(P_1) \cdot f(P_2) < 0$  con  $f(P_1)$  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 uno zero", ossia intersecare il piano  $z = 0$ .

## 1.2 Limiti e continuità

### 69. Come si definisce una il limite di una funzione a più variabili attraverso gli intorni?

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

**Definizione 1.2.1 (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  $f(\underline{x}) \rightarrow l$ , ovvero che tende a  $l$ , se  $\underline{x} \rightarrow \underline{x}_0$ , con  $l \in \mathbb{R}^* = \mathbb{R} \cup \{\pm\infty\}$

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

Visto che  $l \in \mathbb{R}^*$  si parla di intorni di infinito, ossia semirette del tipo  $(-\infty, n)$  e  $(n, +\infty)$ . Quindi,  $\forall U \subset \mathbb{R}^*$  di  $l$ , esiste un intorno sferico di  $\underline{x}_0$ :

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

**Definizione 1.2.2 ( $\epsilon - \delta$  con  $l$  finito).** Se  $\lim_{\underline{x} \rightarrow \underline{x}_0} f(\underline{x}) = l$  ed  $l$  è finito, allora

$$\begin{aligned} \forall \epsilon > 0, \exists \delta > 0 \text{ tale che } |f(\underline{x} - l)| < \epsilon \\ \forall \underline{x} \in (A - \{\underline{x}_0\}) \text{ con } \|\underline{x} - \underline{x}_0\| < \delta = 0 < \|\underline{x} - \underline{x}_0\| < \delta \text{ e } \underline{x} \in B(\underline{x}_0, \delta) \end{aligned}$$

$|f(\underline{x} - l)|$  equivale a  $l - \epsilon < f(\underline{x}) < l + \epsilon = B(l, \epsilon)$  con  $f(\underline{x}) \in B(l, \epsilon)$

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

**Definizione 1.2.3 ( $\epsilon - \delta$  con  $l = \pm\infty$ ).** Se  $\lim_{\underline{x} \rightarrow \underline{x}_0} f(\underline{x}) = \pm\infty$ , quindi il limite diverge, allora

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

Penso 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  $\infty$ .

## 70. Quali sono le proprietà fondamentali del limite in più variabili?

Le proprietà fondamentali del limite:

### 1. Unicità

Se esiste, è unico. Informalmente, se due limiti della stessa funzione multivariata esistono finiti e danno lo stesso risultato vuol dire che il limite è necessariamente unico. Se  $\lim_{\underline{x}_1 \rightarrow \underline{x}_0} f(\underline{x}_1) = l_1$  e  $\lim_{\underline{x}_2 \rightarrow \underline{x}_0} f(\underline{x}_2) = l_2$  esistono finiti e  $l_1 = l_2$  allora il limite esiste ed è unico.

### 2. 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)

## 71. Quando è un limite ben definito?

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.

Quindi, se scelti due cammini e i limiti risultano diversi allora il limite per  $f(x, y)$  non esiste.

## 71. Cosa si intende per limiti ristretti a sottoinsiemi?

Se  $\lim_{(x,y) \rightarrow (x_0,y_0)} f(x, y) = l$  in  $A$ , allora per ogni sottoinsieme  $C \subseteq A$  si ha:

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

In altre parole, se il limite esiste nell'insieme  $A$ , allora esiste ed è uguale a  $l$  anche quando ci si restringe a qualsiasi sottoinsieme  $C$  di  $A$ .

Ovviamente, il punto  $(x_0, y_0)$  deve essere un **punto di accumulazione** per il dominio  $A$ .

Formalmente:  $(x_0, y_0) \in \overline{A} \setminus \{(x_0, y_0)\}$  oppure in ogni intorno di  $(x_0, y_0)$  esistono infiniti punti di  $A$  distinti da  $(x_0, y_0)$ .

Il punto  $(x_0, y_0)$  deve essere un punto di accumulazione anche per il sottoinsieme  $C$ .

Formalmente:  $(x_0, y_0) \in \overline{C} \setminus \{(x_0, y_0)\}$ .

## 72. Come si dimostra la non esistenza di un limite?

Se esistono due sottoinsiemi  $C_1, C_2 \subseteq A$  tali che  $(x_0, y_0)$  è punto di accumulazione per entrambi, e:

$$\lim_{\substack{(x,y) \rightarrow (x_0,y_0) \\ (x,y) \in C_1}} f(x, y) = l_1 \neq l_2 = \lim_{\substack{(x,y) \rightarrow (x_0,y_0) \\ (x,y) \in C_2}} f(x, y)$$

allora il limite  $\lim_{(x,y) \rightarrow (x_0,y_0)} f(x, y)$  non esiste.

Se il limite lungo due curve diverse che passano per  $(x_0, y_0)$  dà risultati diversi, allora il limite della funzione per  $(x, y) \rightarrow (x_0, y_0)$  non esiste. Questo è una diretta conseguenza del teorema di unicità del limite.

E' possibile osservare che come condizione necessaria per l'esistenza del limite è spesso utilizzato il fatto che i limiti, lungo le rette parallele agli assi coordinati di equazione  $y, x = k$  con  $k \in \mathbb{R}$  (costanti), quando esistono sono uguali fra loro, ottenendo come condizione necessaria:

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

### 1.2.1 Continuità

#### 73. Come si definisce la continuità di un limite su un dominio?

**Definizione 1.2.4 ( $\epsilon - \delta$  con la continuità).** Siano  $f(x, y) = \underline{x}$ , quindi  $f : A = \mathbb{R}^2 \rightarrow \mathbb{R}$  (ovvero  $f$  è funzione continua in  $A = \mathbb{R}^2$ ). Per dimostrarlo devo mostrare che sia continua in ogni punto  $\forall P_0 = (x_0, y_0) \in \mathbb{R}^2$ , dobbiamo dimostrare che  $\lim_{x,y \rightarrow x_0, y_0} f(x, y) = f(x_0, y_0)$ .

#### DIMOSTRAZIONE

Si dimostra che  $\forall \epsilon > 0$ ,  $\exists \delta = \delta(\epsilon)$  tale che  $|f(x, y) - f(x_0, y_0)| < \epsilon$ .

Sia  $f(x, y) = \underline{x}$ , vogliamo dimostrare che  $f$  è continua in ogni punto  $P_0 = (x_0, y_0) \in \mathbb{R}^2$ .

Fissato  $\epsilon > 0$ , dobbiamo trovare un  $\delta > 0$  tale che la distanza tra le immagini sia minore di  $\epsilon$ . Osserviamo che:

$$|f(x, y) - f(x_0, y_0)| = |\underline{x} - \underline{x}_0|$$

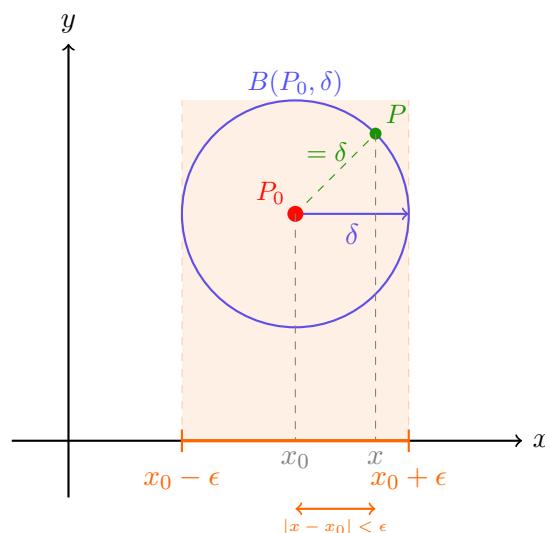
Quindi, possiamo scrivere che la distanza (con proprietà della positività indotta dalla norma):

$$0 \leq |\underline{x} - \underline{x}_0| = \sqrt{(\underline{x} - \underline{x}_0)^2} \leq \sqrt{(x - x_0)^2 + (y - y_0)^2}$$

Poiché la distanza euclidea  $d(P, P_0) = \sqrt{(x - x_0)^2 + (y - y_0)^2}$ , se poniamo  $\delta = \epsilon$ , otteniamo:

$$d(P, P_0) < \delta \implies |\underline{x} - \underline{x}_0| < \delta = \epsilon$$

La condizione è soddisfatta, pertanto la funzione è continua in  $\mathbb{R}^2$ . □



Se  $\delta = \epsilon$ , allora ogni punto dentro il cerchio ha  $|x - x_0| < \epsilon$

La continuità della proiezione  $f(x, y) = \underline{x}$  esprime un fatto geometrico elementare: in un triangolo rettangolo, la lunghezza di un cateto ( $|\underline{x} - \underline{x}_0|$ ) non può mai superare quella dell'ipotenusa ( $d(P, P_0)$ ). Quindi, se limitiamo il movimento totale nel piano entro un raggio  $\delta$ , limitiamo automaticamente la variazione della funzione entro lo stesso valore. Non essendoci "salti" bruschi nelle coordinate mentre ci si sposta nel piano, la funzione è continua.

#### 74. Quando le funzioni sono continue?

**Teorema 1.2.1.** 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.

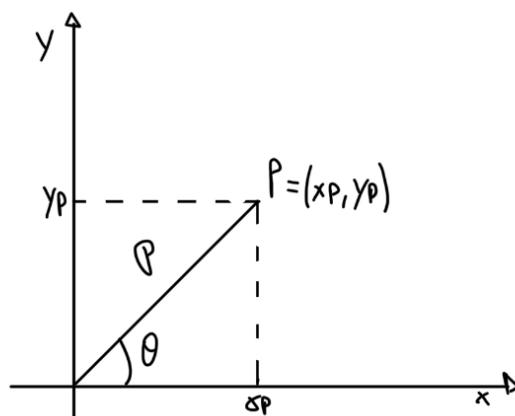
Alcuni esempi:

- $f(x, y) = \sin(x^2 + y^2)$  — 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  $\{(x, y) : x > 0\}$

### 1.2.2 Coordinate Polari

#### 75. Cosa sono le coordinate polari?

**Definizione 1.2.5 (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  $\theta$  formato da una semiretta fissa uscente da  $O$  (detta asse polare) e il segmento che unisce  $O$  a  $P$ , misurato in senso antiorario.



Abbiamo specificato cos'è la distanza  $\rho$  ma è **importante** e necessario spiegare  $\theta$ : geometricamente l'angolo  $\theta$  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  $\theta$  delle coordinate polari.

$$\tan \theta = \frac{\sin \theta}{\cos \theta} \text{ in coordinate polari} \Rightarrow \tan \theta = \frac{\rho \sin \theta}{\rho \cos \theta} = \frac{y_\rho}{x_\rho}$$

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

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

### 76. Come si studiano i limiti con le coordinate polari?

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,  $\rho$  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  $\theta$ , ossia il valore non dipenda dalla direzione da cui ci si avvicina a  $(x_0, y_0)$ .

### 77. Perché è importante che il limite sia uniforme rispetto all'angolo $\theta$ ?

Come scritto sopra, il valore del limite **non** deve dipendere dalla direzione da cui ci si avvicina.

#### Definizione 1.2.6 ( $\epsilon - \delta$ ).

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

$\theta$  rappresenta tutte le possibili direzioni da cui ci si avvicina, se il limite dipende da  $\theta$  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  $\theta$ , violando l'uniformità rispetto a  $\theta$  e affermando che il limite non esiste.

### 78. Come si può usare il teorema del confronto per dimostrare l'esistenza di un limite di una funzione multivariata?

**Definizione 1.2.7 (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:

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

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

- **Indipendenza da  $\theta$  (Funzione Radiale):**  $g$  deve dipendere esclusivamente dalla distanza  $\rho$ . 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  $\theta$ ;
- **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 (o dei Carabinieri)**.

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

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

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  $\theta$  con cui ci si avvicina al punto  $(x_0, y_0)$ .

### 79. Come si studia la continuità di una funzione in due variabili?

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  $\theta$  verificando l'uguaglianza.

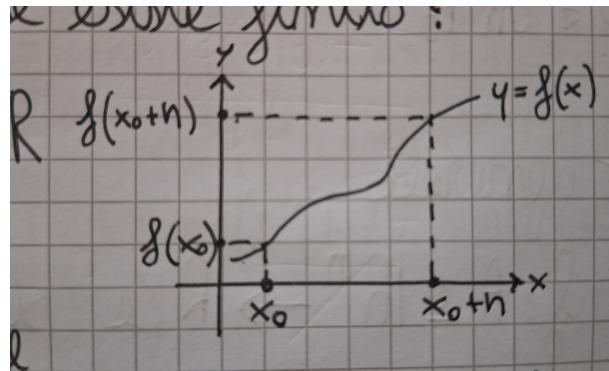
### 1.3 Derivata

**80. Per la derivabilità e differenziabilità, quali sono le principali differenze concettuali rispetto al caso unidimensionale?**

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  $P = (x, f(x))$ .



Nel caso delle funzioni multivariate le direzioni degli assi coordinati ci permettono di definire le derivate parziali di  $f$  (il numero di derivate parziali è dato dal numero di variabili):

- Esistono infinite direzioni lungo cui la funzione può variare;
- Le derivate parziali non sono più sufficienti a garantire la differenziabilità;
- Una funzione può avere tutte le derivate parziali in un punto senza essere continua o differenziabile in quel punto. Serve una condizione più forte: l'esistenza di un'approssimazione lineare con resto;
- Il concetto di gradiente sarà importante per comprendere le direzioni di variazione di una funzione.

#### 1.3.1 Derivata parziale

**81. Cos'è una derivata parziale?**

**Definizione 1.3.1** (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 coordinato 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  $x$  incrementa)

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

- La derivata rispetto a  $y$  (solo  $y$  incrementa)

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

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 1.3.2 (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:

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

### 1.3.2 Gradiente

**82. Quali condizioni devono essere soddisfatte perché una funzione sia derivabile in un punto e su un aperto?**

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  $\nabla f(x_0, y_0) = (\frac{\partial f}{\partial x}(x_0, y_0), \frac{\partial f}{\partial y}(x_0, y_0))$  (nel caso generale  $\nabla f(\underline{x}) = (\frac{\partial f}{\partial x_1}(\underline{x}), \dots, \frac{\partial f}{\partial x_n}(\underline{x}))$ )

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

**83. Perché si parla di vettori e cos'è il vettore gradiente?**

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 1.3.3 (Gradiente).** E' 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  $f(\underline{x} + h) \approx f(\underline{x}) + \langle \nabla f(\underline{x}), \underline{h} \rangle + o(h)$

- **In una variabile**, vicino ad un punto  $x_0$  vale l'approssimazione lineare  $f(x_0 + h) \approx f(x_0) + f'(x_0)h$  con  $h$  una piccola di  $x$ ,  $f'(x_0)$  la pendenza della tangente e  $f'(x_0)h$  di quanto cambia  $f$  quando  $x$  cambia di  $h$ .
- **In più variabili** il punto non si tratta più di un numero ma di un vettore  $\underline{x} = (x_0, x_1, \dots, x_n)$  e anche la variazione non è più un numero ma un vettore  $\underline{h} = (h_1, \dots, h_n)$ .

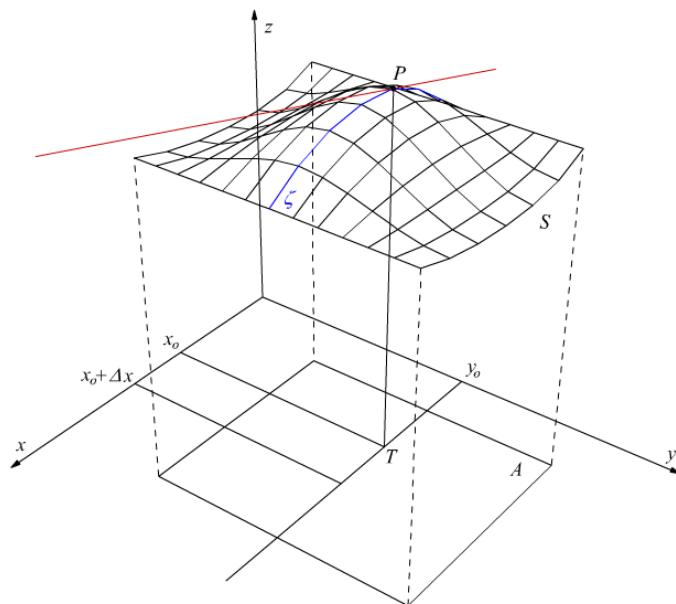
Si definisce il prodotto scalare perché quando ti sposti da  $\underline{x}$  a  $\underline{x} + h$  stai cambiando tutte le variabili della funzione insieme, ossia sommando il contributo di ogni variabile per la corrispettiva derivata parziale

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

Quindi,  $f(\underline{x} + h) \approx f(\underline{x}) + \langle \nabla f(\underline{x}), \underline{h} \rangle$  vicino  $\underline{x}$  la funzione si comporta come un piano tangente, il termine con il prodotto scalare approssima linearmente nel modo migliore e l'errore diventa trascurabile quando  $h$  è piccolo.

#### 84. Come si rappresenta geometricamente una derivata parziale?

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  $yz$ .



#### 85. La derivata parziale è un concetto particolare della derivata direzionale, cosa si intende per derivata parziale in un punto usando la notazione vettoriale?

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

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

- $he_i = h(0\dots1\dots0) = (0\dots h \dots 0)$

- $\underline{x} + h\underline{e}_i = (x_1 \dots x_i + h \dots x_n)$

- $f$  dipende da  $h$

La derivata parziale può essere rappresentata con un concetto più generale: la **derivata direzionale**. Dato l'iperpiano 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 1.3.4 (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*

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

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  $v = \frac{\underline{v}}{\|\underline{v}\|}$ .

## 1.4 Differenziabilità

### 86. Qual è la differenza tra differenziabilità e derivabilità?

- **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 qualcosa di più forte: la funzione può essere sostituita localmente con la sua 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  $x - x_0$ .

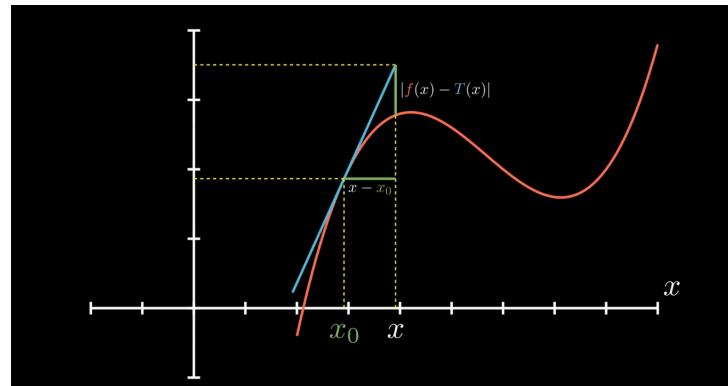
Se dividiamo entrambi i membri per  $x - x_0$  e prendiamo il limite per  $x \rightarrow x_0$ , otteniamo:

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

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):**

A differenza del caso unidimensionale, in più variabili ( $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ) i concetti di **derivabilità** (esistenza delle derivate parziali) e **differenziabilità** si separano nettamente.

La semplice esistenza delle derivate parziali è una condizione debole: garantisce l'esistenza delle rette tangenti solo lungo direzioni specifiche (gli assi), ma non implica la continuità della funzione né l'esistenza di un piano tangente coerente.

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

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

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 condizione di differenziabilità richiede che il termine di errore tenda a zero più velocemente della distanza euclidea:

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

Geometricamente, il valore della funzione vicino a  $\underline{x}_0$  è composto da tre parti:

- Quota base:  $f(\underline{x}_0)$ .
- Variazione sul Piano Tangente:  $\nabla f(\underline{x}_0) \cdot \underline{h}$ .
- Errore di approssimazione:  $o(\|\underline{h}\|)$ , che deve essere trascurabile.

Le derivate parziali forniscono informazioni solo lungo  $n$  direzioni privilegiate ("a croce"). Senza differenziabilità, queste informazioni locali potrebbero non "assemblarsi" in un piano. Potremmo avere una "raggiera" di rette tangenti che non sono complanari.

**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:

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.

### 87. Quindi, come si definisce una funzione differenziabile?

**Definizione 1.4.1 (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:

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

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.

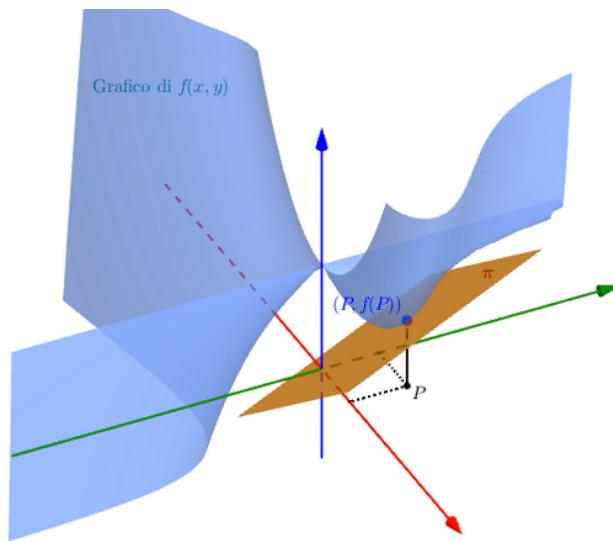
Affermare che il limite è 0 significa che il numeratore (l'errore dell'approssimazione lineare) tende a zero con un **ordine di infinitesimo superiore** rispetto al denominatore (la distanza percorsa  $\|\underline{h}\|$ ). In simboli:

$$f(\underline{x}_0 + \underline{h}) - f(\underline{x}_0) - \nabla f(\underline{x}_0) \cdot \underline{h} = o(\|\underline{h}\|)$$

### 88. Perché la differenziabilità in un punto è legata all'esistenza del piano tangente al grafico della funzione nel punto?

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 iperpiano 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 "zoommiamo" 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.



### 89. Come si ottiene l'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:

$$z = f(x_0, y_0) + \nabla f(P_0) \cdot \underline{h}$$

Dove  $\underline{h} = (x - x_0, y - y_0)$  è il vettore spostamento. Esplicitando in componenti:

$$z = \underbrace{f(x_0, y_0)}_{\text{punto di contatto}} + \underbrace{\frac{\partial f}{\partial x}(P_0)(x - x_0) + \frac{\partial f}{\partial y}(P_0)(y - y_0)}_{\text{inclinazione}}$$

Questa formula è la naturale estensione geometrica della retta tangente. In dimensione 1, l'equazione della retta tangente è:

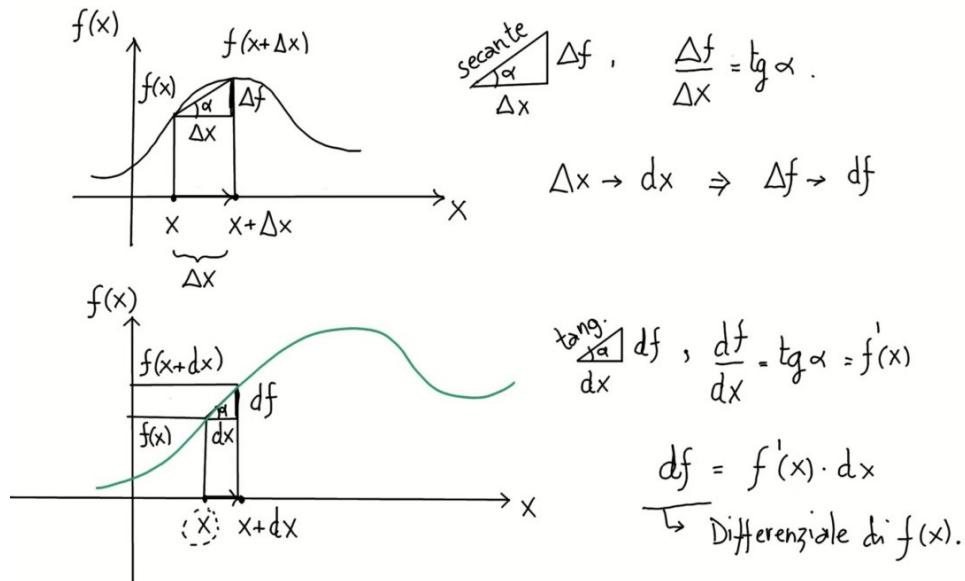
$$y = f(x_0) + f'(x_0)(x - x_0)$$

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.

### 90. Qual è il significato geometrico del differenziale?



Nell'analisi multivariata, dato un punto base  $\underline{x}_0$  e un vettore spostamento  $\underline{h}$ :

- **L'incremento vero ( $\Delta f$ ):** 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  $\Delta f$  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 ( $o(\|\underline{h}\|)$ ).

### 91. Se è differenziabile allora è continua?

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:

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

Passando al limite per  $\underline{h} \rightarrow \underline{0}$  in entrambi i membri:

$$\lim_{\underline{h} \rightarrow \underline{0}} f(\underline{x}_0 + \underline{h}) = \lim_{\underline{h} \rightarrow \underline{0}} [f(\underline{x}_0) + \nabla f(\underline{x}_0) \cdot \underline{h} + o(\|\underline{h}\|)]$$

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  $\underline{x}_0$ . □

### 92. Enunciare e dimostrare il teorema del differenziale.

**Teorema 1.4.1 (Teorema del differenziale).** *Sia  $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  con  $A$  aperto e si suppone che  $f$  sia derivabile su tutto  $A$  tale che*

$$\exists \nabla f(\underline{x}), \forall \underline{x} \in A \Leftrightarrow \text{esistono } f_{x_1}(\underline{x}), f_{x_2}(\underline{x}), \dots, f_{x_n}(\underline{x})$$

*Se le  $n$  derivate sono continue allora  $f$  è differenziabile nel punto  $\underline{x}$ .*

**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  $\underline{h} = (h, k)$

$$\lim_{h, k \rightarrow 0} \frac{f(x + h, y + k) - f(x, y) - \nabla f(x, y) \cdot (h, k)}{\sqrt{h^2 + k^2}} = 0$$

**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 di Lagrange

Poiché  $f$  è derivabile, applichiamo il Teorema di Lagrange (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  $\Delta f$  con l'espressione trovata con Lagrange. 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 diseguaglianza 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  $\leq 1$  (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. □

### 93. Come si può definire una gerarchia di regolarità (funzioni di classe $C^k$ )?

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).

### 94. Se la funzione è differenziabile in un punto, quindi continua, ammette tutte le derivate direzionali nel punto?

Sia  $f : A \rightarrow \mathbb{R}$ , con  $A \subset \mathbb{R}^n$  aperto e  $\underline{x}_0 \in A$ . 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'ō-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. □

**95. Perché diciamo che il vettore **gradiente** indica la direzione e il verso della massima pendenza di una funzione multivariata in un punto?**

Per comprendere questo fatto geometrico, usiamo la **regola della catena**. Consideriamo una curva regolare che passa per il punto  $\underline{x}_0$ :

$$\gamma : [-1, 1] \subset \mathbb{R} \rightarrow A \subseteq \mathbb{R}^n$$

Tale che  $\gamma(0) = \underline{x}_0$  e  $\gamma'(0) = \underline{v}$  (dove  $\underline{v}$  è il vettore direzione unitario,  $\|\underline{v}\| = 1$ ).

Definiamo la funzione composta (restrizione di  $f$  alla curva):

$$F(t) = (f \circ \gamma)(t) = f(\gamma(t))$$

Questa funzione rappresenta la "quota" percorrendo la curva  $\gamma$ . Poiché  $f$  è differenziabile, la derivata di  $F$  in  $t = 0$  (la variazione di quota istantanea) è data dalla regola della catena:

$$F'(0) = \nabla f(\underline{x}_0) \cdot \gamma'(0) = \nabla f(\underline{x}_0) \cdot \underline{v}$$

Informalmente:

- $\underline{x}_0$  è il punto dove ci si trova sulla superficie;
- la curva  $\gamma(t)$  è il "sentiero" che si decide di percorrere;
- $\gamma'(0) = \underline{v}$  è il vettore che rappresenta verso dove ci stiamo muovendo;
- $F(t) = f(\gamma(t))$  è la quota (altitudine) a cui si cammina.

La derivata  $F'(0)$  misura quanto velocemente cambia la quota nel tempo, ossia la pendenza sul "sentiero". La regola della catena ci dice praticamente che questa pendenza dipende dal prodotto scalare  $\nabla f(\underline{x}_0) \cdot \underline{v}$ .

Ora, immaginando di scegliere una direzione qualsiasi  $\underline{v}$  in cui muoverci ma con il vincolo che la velocità sia unitaria ( $\|\underline{v}\| = 1$ ), possiamo comprendere quale direzione ci fa salire più velocemente.

Secondo la proposizione precedente, la derivata direzionale  $D_{\underline{v}}f$  è data da:

$$D_{\underline{v}}f(\underline{x}_0) = \nabla f(\underline{x}_0) \cdot \underline{v}$$

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  $\theta$  è l'angolo tra il gradiente e la direzione scelta  $\underline{v}$ .

Per massimizzare la pendenza  $D_{\underline{v}}f$ , dobbiamo massimizzare  $\cos \theta$ . Il massimo si ha per  $\theta = 0$ , ovvero quando la direzione  $\underline{v}$  è **parallela e concorde** al gradiente  $\nabla f(\underline{x}_0)$ .

Analizzando il termine  $\cos \theta$ :

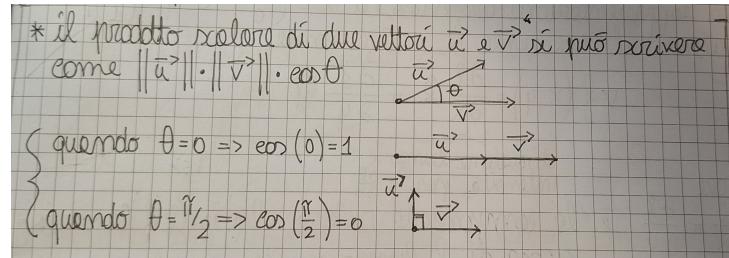
- **Massima Crescita ( $\theta = 0$ )**: La derivata è massima quando ci muoviamo nella direzione del gradiente.

$$D_{\underline{v}}f = \|\nabla f\|$$

- **Massima Decrescita ( $\theta = \pi$ ):** La derivata è minima (negativa) quando ci muoviamo in direzione opposta al gradiente.

$$D_{\vec{v}} f = -\|\nabla f\|$$

- **Crescita Nulla ( $\theta = \pi/2$ ):** La funzione non cambia valore se ci muoviamo ortogonalmente al gradiente (lungo le curve di livello).



### 96. Cosa si intende per **formula del gradiente?**

Sia  $f$  differenziabile su  $A$  e sia un vettore  $\vec{v}$  normato, ossia  $\|\vec{v}\| = \sqrt{v_1^2 + v_2^2}$ , si può definire la derivata direzionale con la formula

$$D_{\vec{v}} f(x, y) = \langle f(x, y), \vec{v} \rangle$$

### 97. Enunciare e dimostrare la regola della derivazione composta o **regola della catena**.

Se  $f$  è differenziabile e  $\gamma$  è derivabile, allora la derivata della funzione composta  $F(t) = f(\gamma(t))$  è:

$$F'(t) = \langle \nabla f(\gamma(t)), \gamma'(t) \rangle$$

Che equivale al prodotto scalare tra il gradiente di  $f$  e il vettore tangente alla curva  $\gamma$ .

### DIMOSTRAZIONE

Consideriamo il rapporto incrementale di  $F(t)$  rispetto a un incremento  $h \in \mathbb{R}$ :

$$\frac{F(t+h) - F(t)}{h} = \frac{f(\gamma(t+h)) - f(\gamma(t))}{h}$$

Poiché per ipotesi  $f$  è differenziabile nel punto  $x = \gamma(t)$ , 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:

$$= \left\langle \nabla f(\gamma(t)), \frac{\gamma(t+h) - \gamma(t)}{h} \right\rangle + \frac{o(\|\gamma(t+h) - \gamma(t)\|)}{h}$$

Passiamo ora al limite per  $h \rightarrow 0$ :

Poiché  $\gamma$  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. Dobbiamo ora dimostrare che il resto tende a 0.

Dobbiamo mostrare che:

$$\lim_{h \rightarrow 0} \frac{o(\|\gamma(t+h) - \gamma(t)\|)}{h} = 0$$

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  $\gamma$ , quando  $h \rightarrow 0$  anche l'incremento  $\|\gamma(t+h) - \gamma(t)\| \rightarrow 0$ .
- **Fattore (B):**

$$\lim_{h \rightarrow 0} \left[ \sum_{i=1}^n \left( \frac{\gamma_i(t+h) - \gamma_i(t)}{h} \right)^2 \right]^{\frac{1}{2}} = \left[ \sum_{i=1}^n (\gamma'(t))^2 \right]^{\frac{1}{2}} \in \mathbb{R}$$

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:

$$F'(t) = \langle \nabla f(\gamma(t)), \gamma'(t) \rangle + 0$$

□

## 1.5 Derivate successive

### 98. Cosa sono le **derivate successive** di una funzione multivariata?

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.

### 1.5.1 Matrice Hessiana

#### 99. Qual è l'equivalente del gradiente per le derivate seconde di una funzione?

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

$$f_{x_i x_j}(\underline{x}) = \frac{\partial^2 f}{\partial x_i \partial x_j}(\underline{x})$$

#### 100. Come si definisce una 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  $2 \times 2$

$$Hf(x, y) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2}(x, y) & \frac{\partial^2 f}{\partial x \partial y}(x, y) \\ \frac{\partial^2 f}{\partial y \partial x}(x, y) & \frac{\partial^2 f}{\partial y^2}(x, y) \end{bmatrix}$$

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.

#### 101. Enunciare e dimostrare il teorema di Schwarz.

**Teorema 1.5.1 (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  $D_{x_i x_j}(\underline{x}) \quad \forall i, j = 1 \dots n$  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  $f$ :

$$F(x) = f(x, y) - f(x, y_0) \quad (\text{con } \forall y \text{ fissato})$$

$$G(y) = f(x, y) - f(x_0, y) \quad (\text{con } \forall x \text{ fissato})$$

**Passo 1:** Applicando il teorema di Lagrange 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 Lagrange 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 Lagrange 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 controsensi (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  $(0, 0)$ .

## 1.6 La formula di Taylor in funzioni multivariate

### 102. Come si può generalizzare la formula di Taylor a più dimensioni?

Dire che  $f \in C^2(A)$  significa che la matrice hessiana  $Hf(A)$  è simmetrica in ogni punto dell'aperto  $A$ :

$$Hf(\underline{x}) = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix} \quad \text{è una matrice simmetrica.}$$

Questo implica che  $Hf = Hf^T$ .

In generale, si dice che  $f \in C^k(A)$ , con  $k \in \mathbb{N}$ , se tutte le derivate parziali fino all'ordine  $k$  esistono e sono continue.

L'*idea fondamentale* si basa sull'analizzare il comportamento di una funzione  $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  che si suppone sia di classe  $C^2(A)$  lungo un segmento rettilineo.

Siano  $\underline{x}, \underline{x} + \underline{h} \in A \subseteq \mathbb{R}^n$  con  $\underline{h} \neq 0$ .

Poiché  $A$  è aperto, per  $\underline{h}$  abbastanza piccolo possiamo dire che il segmento di estremi  $\underline{x}$  e  $\underline{x} + \underline{h}$  è 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:

$$\underline{x}(0) = \underline{x}, \quad \underline{x}(1) = \underline{x} + \underline{h}.$$

Considerando una funzione composta (applicando la regola di derivazione delle funzioni composite) 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:

$$F'(t) = (f(\underline{x} + t\underline{h}))' = \langle \nabla f(\underline{x} + t\underline{h}), \underline{x}'(t) \rangle.$$

Poiché  $x'(t) = \underline{h}$ , si ha:

$$F'(t) = \langle \nabla f(\underline{x} + t\underline{h}), \underline{h} \rangle = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\underline{x} + t\underline{h}) h_i.$$

Poiché abbiamo supposto che  $f \in C^2(A)$ , la funzione  $F(t)$  è derivabile due volte. Per calcolare  $F''(t)$  dobbiamo derivare nuovamente l'espressione di  $F'(t)$  rispetto alla variabile  $t$ . 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

$$F(1) = F(0) + F'(0) + \frac{F''(0)}{2!} + \cdots + \underbrace{\frac{F^{k-1}(0)}{k-1}}_{R_n} + \underbrace{\frac{F^k(\theta)}{k!}}$$

*Informalmente*, invece di studiare il comportamento della funzione  $f$  nell'intero dominio multidimensionale  $A \subseteq \mathbb{R}^n$ , stiamo "affettando" il grafico della funzione lungo una retta. Definendo  $F(t) = f(\underline{x} + t\underline{x})$ , 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.

**103. Enunciare e dimostrare la formula di Taylor secondo il resto secondo Lagrange.**

La formula di Taylor rispetto a  $F'$  con resto secondo Lagrange equivale al teorema di Lagrange per funzioni in più variabili.

**Teorema 1.6.1.** Sia  $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  con  $A$  un aperto e  $f \in C^k(A)$ .

Ipotizziamo che  $\underline{x}$  e  $(\underline{x} + \underline{h})$  ( $\underline{h} \neq 0$ ) siano due elementi di  $A$  per cui il segmento di estremi  $[\underline{x}, (\underline{x} + \underline{h})] \subset A$ . Si può trovare un  $\theta$  reale  $\in [0, 1]$  tale che

Nel caso in cui  $k = 1$ :

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

che equivale alla formula di Taylor  $T_n \Rightarrow F(1) = F(0) + F'(0)$

Nel caso in cui  $k = 2$ :

$$f(\underline{x} + \underline{h}) = f(\underline{x}) + \langle \nabla f(\underline{x} + \theta \underline{h}), \underline{h} \rangle + \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(\underline{x} + \theta \underline{h}) h_i h_j$$

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  $n \times n$ :

$$Hf(\underline{x}) \cdot \underline{h} \quad \text{dove } \underline{h} \text{ è un vettore } (n \times 1)$$

In termini matriciali:

$$\underbrace{\begin{bmatrix} \dots \\ \dots \end{bmatrix}}_{n \times n} \cdot \underbrace{\begin{bmatrix} \vdots \\ \vdots \end{bmatrix}}_{n \times 1} \rightarrow \text{Vettore colonna}$$

Esplicitamente la matrice Hessiana è:

$$Hf(\underline{x} + \theta \underline{h}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(\underline{x} + \theta \underline{h}) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(\underline{x} + \theta \underline{h}) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(\underline{x} + \theta \underline{h}) & \dots & \frac{\partial^2 f}{\partial x_n^2}(\underline{x} + \theta \underline{h}) \end{bmatrix}$$

□

**103. Enunciare e dimostrare la formula di Taylor secondo il resto secondo Peano.**

Sia  $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  e  $f \in C^2(A)$ . Siano  $\underline{x}$  e  $\underline{x} + \underline{h}$  (distinti, per  $\underline{h} \neq \underline{0}$ ) appartenenti ad  $A$ , per cui il segmento con estremi  $[\underline{x}, \underline{x} + \underline{h}] \subseteq A$ . 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:

$$\text{Espressione omogenea di } 2^\circ \text{ grado in } h, k : \quad \frac{\partial^2 f}{\partial x^2}h^2 + 2\frac{\partial^2 f}{\partial x \partial y}hk + \frac{\partial^2 f}{\partial y^2}k^2$$

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  $\leq 2$  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  $2^\circ$  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.

## 1.7 Studio dei massimi e minimi di funzioni multivariete

### 104. Come si definiscono massimi e minimi locali di una funzione multivariata?

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 1.7.1 (Massimo e minimo locale).** Se  $\exists B(\underline{x}_0, r)$  con  $r > 0$  (intorno di  $\underline{x}_0$ ) per cui  $\forall x \in A \cap B(\underline{x}_0, r)$   $f(\underline{x}_0) \geq f(\underline{x})$  allora  $\underline{x}_0$  è un punto di minimo locale di  $f$ , mentre se  $f(\underline{x}_0) \leq f(\underline{x})$  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.

### 105. Qual è la differenza tra ottimizzazione vincolata e 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. Invece, l'ottimizzazione vincolata implica la ricerca di massimi e minimi di una funzione soggetta a determinate restrizioni o vincoli, di solito espressi come equazioni o disequazioni.

## 1.8 Ottimizzazione libera

### 106. Enunciare e dimostrare il teorema di Fermat per funzioni multivariate.

**Teorema 1.8.1 (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 massimo (minimo) 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

Sia  $\underline{x}_0 \in A$  un punto di estremo relativo per  $f$ , ad esempio un massimo locale, allora  $\underline{x}_0$  è un punto di estremo relativo per  $f$  lungo una qualsiasi retta passante per il punto. Consideriamo una direzione arbitraria data da un vettore unitario  $\underline{v} \in \mathbb{R}^n$  e definiamo la funzione ausiliaria  $F(t) = f(\underline{x}_0 + t\underline{v})$  in un intorno di  $t = 0$ , ovvero con  $\delta > 0$  tale che  $B(\underline{x}_0, \delta) \subset A$ .

Tale funzione reale di una variabile  $F(t)$  con  $t \in (-\delta, \delta)$  ha in  $t = 0$  un punto di massimo locale per cui, per il teorema di Fermat, si ha:

$$F'(0) = \langle \nabla f(\underline{x}_0), \underline{v} \rangle = 0$$

□

### 1.8.1 Test dell'hessiana

### 107. Come si studia la natura dei punti critici?

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, governato 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.

### 1.8.2 Forme Quadratiche

#### 108. Come possiamo dimostrare quanto detto sul test dell'hessiana?

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:

$$f(\underline{x}_0 + \underline{h}) - f(\underline{x}_0) \approx \frac{1}{2} \underbrace{\langle Hf(\underline{x}_0) \underline{h}, \underline{h} \rangle}_{q(\underline{h})}$$

Il termine  $q(\underline{h})$  è una forma quadratica.

#### 109. Cos'è una forma quadratica?

**Definizione 1.8.1 (Forma Quadratica).** Una forma quadratica in  $\underline{h} \in \mathbb{R}^n$  è un polinomio omogeneo di secondo grado nelle variabili  $h_1, \dots, h_n$ . Ad essa  $q(\underline{h})$  è associata una matrice simmetrica  $A$  tale che:

$$q(\underline{h}) = \langle A\underline{h}, \underline{h} \rangle = \sum_{i,j=1}^n a_{ij} h_i h_j$$

Data la simmetria ( $a_{ij} = a_{ji}$ ), i coefficienti del polinomio corrispondono agli elementi della matrice (i termini misti si dividono per 2).

Dato il polinomio omogeneo:

$$q(\underline{h}) = 3h_1^2 + 10h_1h_2 + 27h_2^2$$

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  $a_{21}$ .

#### 110. Come si classifica una forma quadratica?

Per determinare la natura del punto critico, dobbiamo studiare il segno di  $q(\underline{h})$  per  $\underline{h} \neq \underline{0}$ . Sia  $A$  la matrice simmetrica reale, essa ammette  $n$  autovalori reali  $\lambda_1, \dots, \lambda_n$ .

- $A$  è definita positiva:  $q(\underline{h}) > 0, \forall \underline{h} \neq \underline{0}$ .
  - $\lambda_i > 0$  per tutti gli  $i$ ;
  - Punto di minimo.
- $A$  è definita negativa:  $q(\underline{h}) < 0, \forall \underline{h} \neq \underline{0}$ .
  - $\lambda_i < 0$  per tutti gli  $i$ ;
  - Punto di massimo.
- $A$  è indefinita:  $q(\underline{h})$  assume sia valori positivi sia negativi..
  - Esistono autovalori discordi (almeno un  $\lambda > 0$  e un  $\lambda < 0$ ).
  - Punto di sella.

#### 111. Come si collegano determinante e traccia della matrice agli autovalori per il test dell'hessiana?

Nel caso di 2 variabili, si può dire che gli autovalori  $\lambda_1, \lambda_2$  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  $2 \times 2$  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  $\implies$  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  $\implies$  autovalori discordi  $\rightarrow$  sella.

### 112. Perché si usano le forme quadratiche?

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 ( $\lambda$ ) rappresentano le curvature lungo le direzioni principali, ovvero quelle direzioni speciali in cui la superficie curva senza torsioni.

- Un  $\lambda$  positivo indica una curvatura verso l'alto (convessità).
- Un  $\lambda$  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  $\lambda_1$  e  $\lambda_2$  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 paraboloide 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

### 1.8.3 Caso dubbio: il test dell'hessiana non funziona

#### 113. Cosa si fa quando il test dell'hessiana fallisce ( $\det H = 0$ )?

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  $\Rightarrow x_1$  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  $\Rightarrow x_1$  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  $\Delta f$  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.

## 1.9 Ottimizzazione vincolata

**114. Enunciare e dimostrare il Teorema di Weierstrass per funzioni multivariete.**  
Se si studia la natura dei punti critici quando l'insieme di definizione è compatto, si può usare il Teorema di Weierstrass.

**Teorema 1.9.1 (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:

$$\exists(x_{\min}, y_{\min}), (x_{\max}, y_{\max}) \in K : f(x_{\min}, y_{\min}) \leq f(x, y) \leq f(x_{\max}, y_{\max}) \quad \forall(x, y) \in K$$

**115. In cosa consiste il problema dell'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*).

$$V = \{(x, y) \in \mathbb{R}^2 : \underbrace{g(x, y)}_{g(x, y) - k = 0} = k\} \subseteq \mathbb{R}^2 \quad \text{con } g \in C^1$$

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

### 1.9.1 Parametrizzazione

**116. Come funziona la parametrizzazione o riduzione a una variabile?**

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.

### 1.9.2 Moltiplicatori di Lagrange

**117. Enunciare il Teorema dei Moltiplicatori di Lagrange.**

Siano  $f, g : A \rightarrow \mathbb{R}^2$  funzioni di classe  $C^1(A)$ , supponendo che  $\underline{x}_0 \in V$  un punto di estremo locale per  $f$  sull'insieme  $V = \{(x, y) \in A : g(x, y) = k\}$ .

Se  $\underline{x}_0$  è un *punto regolare* per il vincolo (ovvero  $\nabla g(x, y) \neq (0, 0)$ ), 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).

### 118. Cos'è la **Funzione Lagrangiana** e come si usa?

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:

$$\nabla \mathcal{L}(x, y, \lambda) = \mathbf{0} \Leftrightarrow \begin{cases} \frac{\partial f}{\partial x} - \lambda g_x = 0 \\ \frac{\partial f}{\partial y} - \lambda g_y = 0 \\ -(g(x, y) - k) = 0 \quad (\text{che è il vincolo stesso}) \end{cases}$$

Le soluzioni  $(x, y, \lambda)$  ci forniscono i candidati  $(x, y)$  da valutare, che vanno poi testati con altri metodi per trovare i massimi e minimi assoluti.

### 119. 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  $x \in [-1, 1]$ ). Sostituendo 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.

## 1.10 Integrazione

### 119. Cosa si intende per integrale doppio?

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  $m_{i,j} = \inf f(x, y)$  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  $M_{i,j} = \sup f(x, y)$  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.

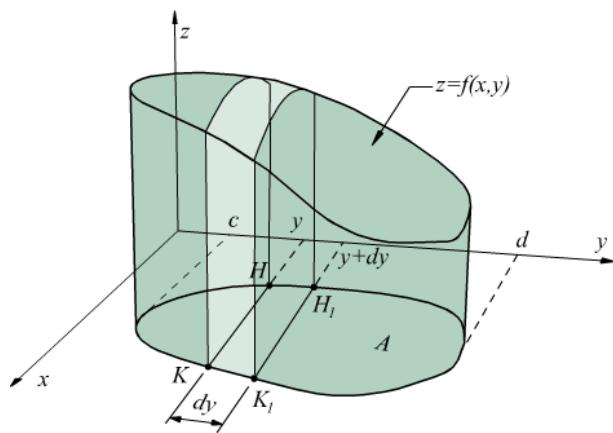
**Definizione 1.10.1 (Integrale Doppio).** Una funzione  $f : R \rightarrow \mathbb{R}$  limitata su  $R$  dove  $R = [a, b] \times [c, d]$ . La funzione  $f$  si dice integrabile secondo Riemann su  $R$  se

$$\sup_D s(f, D) = \inf_D S(f, D)$$

In tal caso, il valore comune (un numero reale) viene detto integrale doppio di  $f$  su  $R$  e si indica con

$$\iint_R f, \quad I(f, R), \quad \iint_R f(x, y) dx dy$$

**120. Qual è l'interpretazione geometrica con cui si definisce l'integrale doppio?** 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 inferiore e superiori è il volume di un parallelepipedo di base il rettangolo  $r_{ij}$  e altezza rispettivamente  $m_{i,j}$  e  $M_{i,j}$ .

### 121. Enunciare il teorema della continuità per gli integrali doppi.

**Teorema 1.10.1 (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  $R$ .*

#### 1.10.1 Calcolo Integrale sui rettangoli

##### 122. Come si calcolano gli integrali doppi su rettangoli? (Formule di riduzione)

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.

##### 123. Come si definisce il teorema di Fubini-Tonelli per gli integrali doppi?

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).

#### 1.10.2 Integrali doppi su domini generali

##### 124. Se invece di rettangoli si avessero dei domini più generali per definire un integrale doppio?

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 1.10.2.** *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:*

$$\iint_D f(x, y) dx dy = \iint_R \tilde{f}(x, y) dx dy$$

**125. Perché non tutti gli insiemi vanno bene e bisogna che insieme sia misurabile?**

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

$$|D| = \iint_D 1 \, dx \, dy \implies \text{area di } D$$

In generale, ogni rettangolo  $R = [a, b] \times [c, d]$  è misurabile secondo Peano-Jordan, e la sua misura è data da:

$$|R| = (b - a)(d - c)$$

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.

**126. C'è 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.

**127. Quali sono le proprietà principali degli 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:

$$\left| \iint_D f(x, y) \, dx \, dy \right| \leq \iint_D |f(x, y)| \, dx \, dy \leq \underbrace{\sup_{(x,y) \in D} |f(x, y)|}_{\text{massimo valore assoluto di } f \text{ in } D} \cdot |D|$$

### 1.10.3 Calcolo Integrale sulle regioni semplici

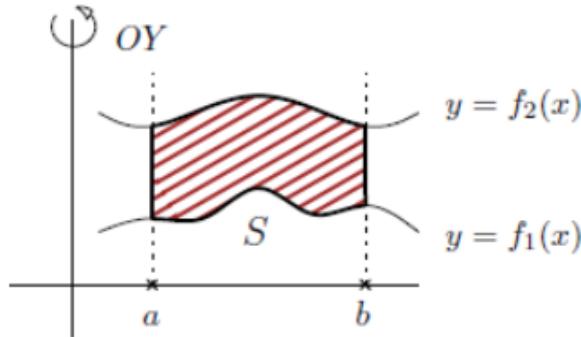
**128. Cosa si intende esattamente per Regione Normale (o Dominio Semplice)?**

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:

$$D = \bigcup_{x \in [a, b]} \{x\} \times [f_1(x), f_2(x)]$$

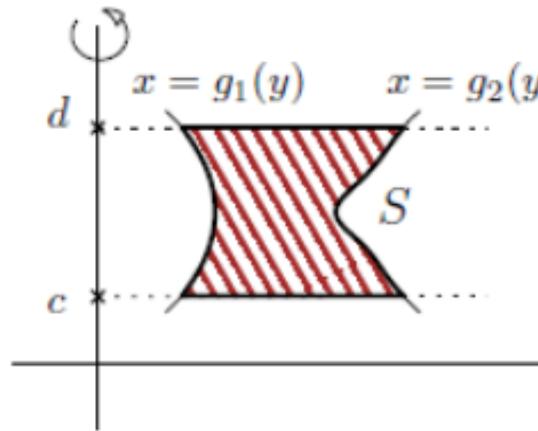
- Dominio normale rispetto all'asse **y (X-simplice)**:

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:

$$D_{y-\text{simplice}} = \int_a^b [f_2(x) - f_1(x)] dx$$

$$D_{x-\text{semplice}} = \int_c^d [h_2(y) - h_1(y)] dy$$

**129.** Quali sono le **formule di riduzione** per questi domini?

**Teorema 1.10.2** (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:

$$\iint_D f(x, y) dx dy = \int_c^d \left( \int_{h_1(y)}^{h_2(y)} f(x, y) dx \right) dy$$

**130.** In generale come si definisce il calcolo gli integrali doppi su **più domini semplici**?

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.

#### 1.10.4 Cambiamento di variabili con integrali doppi

**131.** Perché e quando si effettua un **cambiamento di variabili**?

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 domaini che presentano delle simmetrie ci conviene cambiare le variabili trasformando il dominio  $D$  (nel piano  $xy$ ) in un dominio  $D'$  (nel piano  $\rho, \theta$ ) molto più semplice (ad esempio un rettangolo), tramite una trasformazione biunivoca.

**132. Come si definisce un cambiamento di variabili di una funzione integrabile secondo riemann nell'analisi multivariata?**

Sia  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  (un vettore in un vettore) 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, detta Jacobiana, della  $f(\rho, \theta) \rightarrow (\rho \cos \theta, \rho \sin \theta)$ :

$$J_{f(\rho, \theta)} = \begin{bmatrix} \nabla f_1(\rho, \theta) \\ \nabla f_2(\rho, \theta) \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial \rho}(\rho, \theta) & \frac{\partial f_1}{\partial \theta}(\rho, \theta) \\ \frac{\partial f_2}{\partial \rho}(\rho, \theta) & \frac{\partial f_2}{\partial \theta}(\rho, \theta) \end{bmatrix} = \begin{bmatrix} \cos \theta & -\rho \sin \theta \\ \sin \theta & \rho \cos \theta \end{bmatrix} = \det J_{f(\rho, \theta)} = \rho \cos^2 \theta + \rho \sin^2 \theta = \rho$$

per la proprietà trigonometrica fondamentale  $\cos^2 \theta + \sin^2 \theta = 1$ . Se il determinante è diverso da zero, la trasformazione è invertibile. Con  $\rho = 0 \implies x = 0, y = 0$  l'origine e  $(0, \theta) \rightarrow (0, 0) \forall \theta$

quindi la trasformazione non è biunivoca in  $(0, 0)$ . Possiamo anche affermare che le funzioni trigonometriche sono periodiche ( $2\pi$ ), comportando una non biunivocità con  $(\rho, \theta)$  e  $(\rho, \theta + 2\pi k)$  per ogni intero  $k$ .

Ma se restringiamo il dominio di  $\theta$  a  $[0, 2\pi]$  o  $[-\pi, \pi]$  e  $\rho \geq 0$ , la trasformazione diventa biunivoca. Consideriamo  $f : A \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}^2$  con  $(\rho, \theta) \rightarrow (\rho \cos \theta, \rho \sin \theta)$  e prendiamo  $A = (0, +\infty) \times [0, 2\pi]$  o  $(0, +\infty) \times [-\pi, \pi]$ . Quindi, consideriamo  $(\rho, \theta) \in A$

### 133. Enunciare il teorema del cambiamento di variabili per gli integrali doppi.

**Teorema 1.10.3 (Cambiamento di variabili per integrali doppi).** *Sia  $S$  un sottoinsieme di  $A$  un aperto misurabile secondo Peano-Jordan del piano  $(\rho, \theta)$  e sia  $F : A \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}^2$  una trasformazione di classe  $C^1$  biunivoca tale che  $F(\rho, \theta) = (x(\rho, \theta), y(\rho, \theta))$  con Jacobiana  $J_F(\rho, \theta)$ . Allora per ogni funzione integrabile (limitata e continua) su  $T$  si ha che:*

$$\iint_T f(x, y) \, dx \, dy = \iint_S f(\rho \cos \theta, \rho \sin \theta) \underbrace{\rho}_{|\det J_F(\rho, \theta)|} \, d\rho \, d\theta$$

con  $T = F(S)$  e  $S = F^{-1}(T)$ .

