Corso di Analisi 2

Domande e risposte al macro-argomento: CALCOLO DIFFERENZIALE

Ultima modifica: 1 marzo 2026

Indice

Funzioni di più variabili

1.1 Definizioni e grafici

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

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

Ovvero

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

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

Definizione 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 $\mathbb{R}^2$.

Immagine di una funzione

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

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

1.1.1 Grafico di funzione e sezioni (trasversali e orizzontali)

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

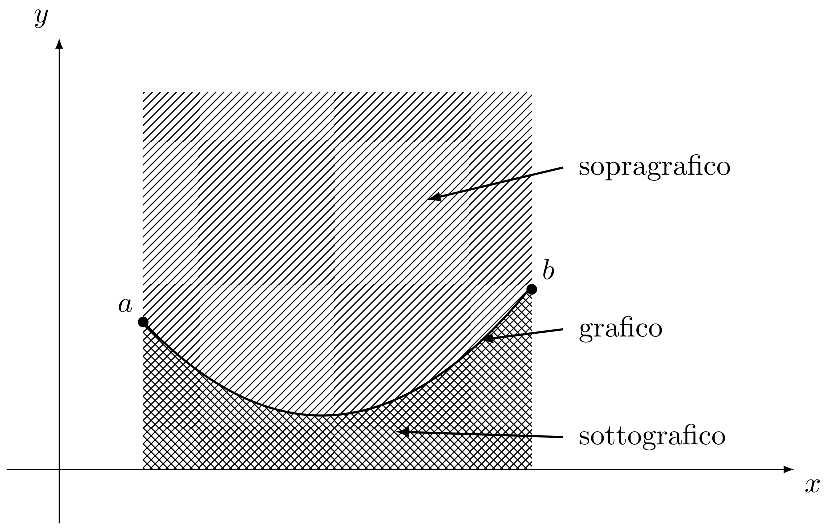
$$\mathrm{graf} f = \{(x, y, z) \in \mathbb{R}^3 : (x, y) \in D, z = f(x, y) \in \mathbb{R}\}$$

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.

Osservazione. Sia $f : I \subseteq \mathbb{R}^2 \to \mathbb{R}$ continua su I e sia il suo grafico $\{(x, y) \in \mathbb{R}^2 : y = f(x)\} \subseteq \mathbb{R}^2$. Il grafico di f descrive non altro che una curva continua sul piano, definita come un arco di curva continua, e in base alla relazione tra y e $f(x)$ potremmo analizzare una sezione del grafico.

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



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

Per visualizzare o rappresentare il grafico di una funzione $f : D \subseteq \mathbb{R}^2 \to \mathbb{R}$ "seziono" il grafico formando le tracce che intersecano il grafico graf$f \subset \mathbb{R}^3$ 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.

Definizione 1.1.4 (Insiemi o curve di livello)**.** *Gli insiemi di livello, o anche curve di livello, di una funzione $f : D \subseteq \mathbb{R}^2 \to \mathbb{R}$ 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

Definizione 1.1.5 (Curva)**.** *Una curva in $\mathbb{R}^n$ è un'applicazione continua definita su un intervallo $I = [a, b] \subseteq \mathbb{R}$*

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

Essa è composta da γ_i componenti definite su $I \subseteq \mathbb{R}$ e continue con $i = 1...n$.
Il sostegno o immagine di una curva è l'insieme dei punti geometrici toccati dalla curva nello spazio $\mathbb{R}^n$

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

Definizione 1.1.6 (Curva Regolare)**.** *Sia una curva $\gamma : I \subseteq \mathbb{R} \to \mathbb{R}^n$ 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)**.** *Una curva cartesiana è una curva che può essere rappresentata come grafico di una funzione $f : I \subseteq \mathbb{R} \to \mathbb{R}$, ossia $\gamma(t) = (t, f(t))$.*

E' il modo di trasformare il grafico di una funzione ordinaria f in una curva parametrica. Percorriamo il grafico come un cammino usando lo stesso x come parametro.
La prima componente scorre sull'asse x e la seconda componente segue il valore della funzione. La curva cartesiana è un caso particolare di curva regolare, infatti se f è derivabile e $f'(t) \neq 0$ allora $\gamma'(t) = (1, f'(t)) \neq 0$.

Definizione 1.1.8 (Curva Chiusa)**.** *Sia una curva $\gamma : I \subseteq \mathbb{R} \to \mathbb{R}^n$, si dice chiusa se il punto iniziale coincide con quello finale $\gamma(a) = \gamma(b)$.*

1.1.3 Accenno al teorema degli zeri per funzioni a più variabili

Teorema 1.1.1 (Teorema degli zeri per funzioni a più variabili)**.** *Sia $f : A \subseteq \mathbb{R}^2 \to \mathbb{R}$ continua su A e supponiamo che esistano due punti $P_1 = (x_1, y_1)$ e $P_2 = (x_2, y_2)$ che appartengono ad un insieme connesso $A_i \subseteq A$ tali che $f(P_1) \cdot f(P_2) < 0$, cioè $f(P_1)$ e $f(P_2)$ hanno segno opposto. Allora esiste almeno un punto $P_0 = (x_0, y_0) \in A_i$ tale che $f(P_0) = 0$.*

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

1.2 Limiti e continuità

1.2.1 I limiti in più variabili e le loro proprietà

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

Definizione 1.2.1 (Limite in $\mathbb{R}^n$). *Sia $\underline{\boldsymbol{x}}_0 = (x_0^1, ..., x_0^n) \in \mathbb{R}^n$ punto di accumulazione per A si dice che $f(\underline{\boldsymbol{x}}) \to l$ se $\underline{\boldsymbol{x}} \to \underline{\boldsymbol{x}}_0$:*

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

Premettiamo che l appartiene all'insieme di numeri reali esteso ($\mathbb{R}^*$) ad intorni di infinito, semirette del tipo $(-\infty, n)$ e $(n, +\infty)$.
Quindi, $\forall U$ intorno di l in $\mathbb{R}^*$, esiste un intorno sferico di $\underline{\boldsymbol{x}}_0$:

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

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

Definizione 1.2.2 ($\epsilon - \delta$ con l finito). *Se*

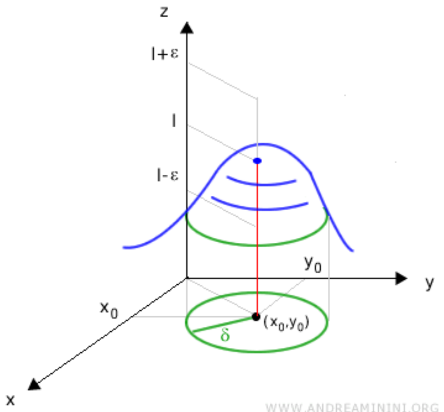
$$\lim_{\underline{\boldsymbol{x}} \to \underline{\boldsymbol{x}}_0} f(\underline{\boldsymbol{x}}) = l$$

allora

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

che equivale a dire

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



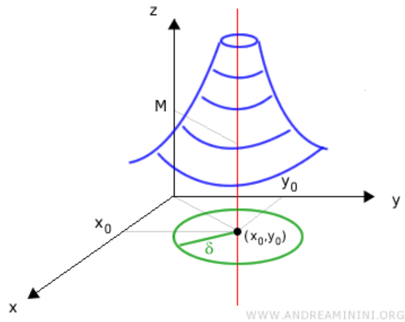
Posso rendere $f(\underline{\boldsymbol{x}})$ vicino a l quando voglio semplicemente prendendo $\underline{\boldsymbol{x}}$ abbastanza vicino a $\underline{\boldsymbol{x}}_0$. In altre parole, più ti avvicini a $\underline{\boldsymbol{x}}_0$ più il valore della funzione si avvicina a l.

Definizione 1.2.3 ($\epsilon - \delta$ con $l = \pm\infty$). *Se*

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

il limite diverge, allora

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



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

- **Unicità**

  Se esiste, è unico.

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

  Se
  $$\lim_{\underline{\boldsymbol{x}}_1 \to \underline{\boldsymbol{x}}_0} f(\underline{\boldsymbol{x}}_1) = l_1 \text{ e } \lim_{\underline{\boldsymbol{x}}_2 \to \underline{\boldsymbol{x}}_0} f(\underline{\boldsymbol{x}}_2) = l_2, \quad \text{con } l_1 = l_2$$
  allora il limite esiste ed è unico.

- **Composizione**

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

Esistenza di un limite

Per esistere un limite deve essere indipendente dalla direzione di avvicinamento di (x, y) al punto di accumulazione $\implies$ un limite non deve dipendere dal cammino o dalla direzione scelti con il quale (x, y) si avvvicina il punto di accumulazione.
Se, scelti due cammini per raggiungere lo stesso punto di accumulazione, il valore dei limiti risulta diverso allora il limite per $f(x, y)$ non esiste.

Definizione 1.2.4 (Limite ristretto ad un sottoinsieme). *Sia $f : A \subset \mathbb{R}^2 \to \mathbb{R}$ e sia (x_0, y_0) punto di accumulazione per A. Se*

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

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

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

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

Sia $f : A \subset \mathbb{R}^2 \to \mathbb{R}$ e sia (x_0, y_0) punto di accumulazione per A.
Se esistono due sottoinsiemi $C_1, C_2 \subseteq A$ tali che:

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

- esistono i limiti ristretti
$$\lim_{\substack{(x,y)\to(x_0,y_0)\\(x,y)\in C_1}} f(x,y) = l_1$$
$$\lim_{\substack{(x,y)\to(x_0,y_0)\\(x,y)\in C_2}} f(x,y) = l_2$$
- con $l_1 \neq l_2$,

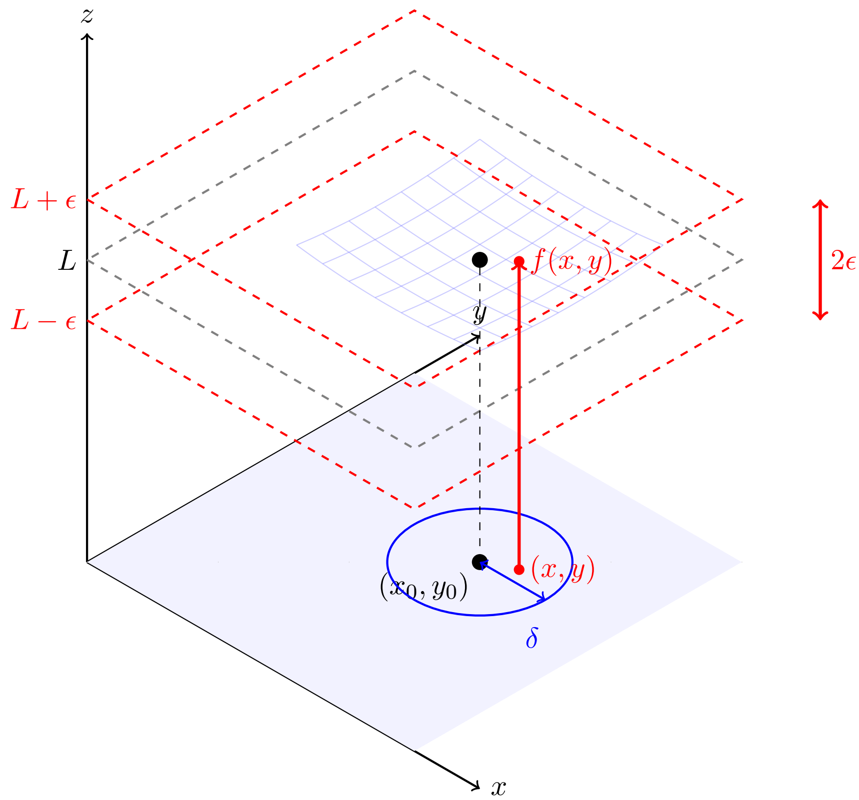
allora
$$\lim_{(x,y)\to(x_0,y_0)} f(x,y) \quad \underline{\text{non esiste}}$$
Se il limite totale esistesse ed fosse uguale a l, allora per ogni sottoinsieme $C \subseteq A$ tale che (x_0, y_0) sia punto di accumulazione per C, dovrebbe valere
$$\lim_{\substack{(x,y)\to(x_0,y_0)\\(x,y)\in C}} f(x,y) = l$$
Questo contraddice l'ipotesi $l_1 \neq l_2$. Il risultato segue dall'unicità del limite.
Come condizione necessaria ma non sufficiente per l'esistenza del limite, si può verificare che i limiti lungo diverse curve (ad esempio rette $y = mx$,oppure curve come $y = x^2$) che tendono a (x_0, y_0) diano lo stesso valore. Se lungo due curve si ottengono due valori limite diversi, il limite della funzione non esiste.

1.2.2 Continuità

Definizione 1.2.5 (Continuità in un punto). *Sia $f : A \subset \mathbb{R}^2 \to \mathbb{R}$. La funzione f si dice continua nel punto $P_0 = (x_0, y_0) \in A$ se*
$$\lim_{(x,y)\to(x_0,y_0)} f(x,y) = f(x_0, y_0)$$
Definizione 1.2.6 (Continuità su un dominio). *La funzione f si dice continua in A se è continua in ogni punto $P_0 \in A$.*

Definizione 1.2.7 ($\epsilon - \delta$ per la continuità). *La funzione f è continua in $P_0 = (x_0, y_0)$ se*
$$\forall \varepsilon > 0,\ \exists \delta > 0 \text{ tale che}$$
$$\|(x,y) - (x_0,y_0)\| < \delta \ \Rightarrow\ |f(x,y) - f(x_0,y_0)| < \varepsilon.$$

DIMOSTRAZIONE
Sia $P_0 = (x_0, y_0)$.
Osserviamo che
$$|f(x,y) - f(x_0,y_0)| = |x - x_0|.$$
Poiché vale la disuguaglianza
$$|x - x_0| \leq \sqrt{(x-x_0)^2 + (y-y_0)^2} = \|(x,y) - (x_0,y_0)\|,$$
se scegliamo $\delta = \varepsilon$, otteniamo:
$$\|(x,y) - (x_0,y_0)\| < \delta \ \Rightarrow\ |x - x_0| < \varepsilon.$$
Quindi f è continua in P_0.
Poiché P_0 è arbitrario, f è continua in tutto $\mathbb{R}^2$. □



Il grafico mostra la relazione tra "vicinanza nel dominio" e "vicinanza nel codominio".
Fissiamo una tolleranza ϵ sul valore della funzione (la banda rossa tra $L-\epsilon$ e $L+\epsilon$). La definizione di limite garantisce che esiste un raggio δ nel piano x-y (il disco blu) tale che: se prendiamo qualunque punto (x, y) dentro questo disco, il corrispondente valore $f(x, y)$ sulla superficie cade necessariamente dentro la banda rossa.
Geometricamente, questo significa che la superficie $z = f(x, y)$ non presenta "salti" o discontinuità in (x_0, y_0): restringendo progressivamente il disco blu (riducendo δ), i valori della funzione si concentrano sempre più attorno a L. Questo è l'essenza della continuità: piccole variazioni in (x, y) producono piccole variazioni in $f(x, y)$.

Teorema 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 e 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 $x_1, x_2, \ldots, x_n$ sono continue. Quindi, costruendo il polinomio a partire dalle coordinate tramite operazioni algebriche (somme, prodotti, multipli), otteniamo sempre una funzione continua.*

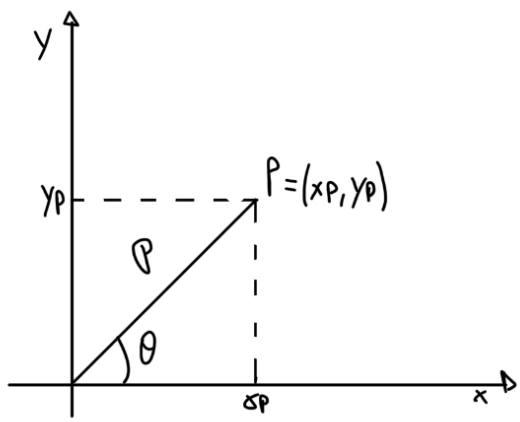
Se abbiamo operazioni tra funzioni continue, il risultato è ancora continuo.
Questo teorema ci permette di affermare che le funzioni elementari, come polinomi a più variabili, funzioni razionali, funzioni esponenziali a più variabili, funzioni trigonometriche, sono tutte continue.

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 $\{(x, y) : x > 0\}$

1.2.3 Coordinate Polari

Definizione 1.2.8 (Coordinate polari). *Un modo alternativo alle coordinate cartesiane per descrivere posizioni sul piano sono le coordinate polari, un sistema bidimensionale che identifica un punto P attraverso due parametri: la distanza $\rho = \sqrt{(x - x_0)^2 + (y - y_0)^2}$ da un punto fisso O (detto polo), e l'angolo θ formato da una semiretta fissa uscente da O (detta asse polare) e il segmento che unisce O a P, misurato in senso antiorario.*



Sappiamo bene il concetto di distanza (ρ), mentre dobbiamo spiegare bene il concetto di angolo (θ).
Geometricamente l'angolo θ è strettamente legato alla coefficente 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 coefficente angolare della retta è $\tan\theta = \frac{\sin\theta}{\cos\theta}$, che va a coincedere esattamente con la definizione di coefficente angolare dell'angolo θ delle coordinate polari:

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

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

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

Per studiare il limite di una funzione $f(x, y)$ quando $(x, y) \to (x_0, y_0)$, 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 $\rho \to 0^+$, perché essendo una radice quadrata di una somma di quadrati, ρ non può mai essere negativa, quindi si avvicina al centro resitingendosi verso (x_0, y_0);

3. Si verifica che il limite sia uniforme rispetto a θ, ossia il valore non dipenda dalla direzione da cui ci si avvicina a (x_0, y_0).

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

Definizione 1.2.9 ($\epsilon - \delta$)**.**

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

θ rappresenta tutte le possibili direzioni da cui ci si avvicina, se il limite dipende da θ significa che per cammini diversi otteniamo valori diversi e per l'esistenza del limite i valori devono essere gli stessi qualsiasi sia la direzione di avvicnamento.
Quindi, se il limite per $f(\rho, \theta) = g(\theta)$ per $\rho \to 0^+$ vuol dire che $g(\theta)$ assume valori diversi al variare di θ, violando l'uniformità rispetto a θ e affermando che il limite non esiste.

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

Definizione 1.2.10 (Il teorema del confronto (o dei due carabinieri))**.** *Siano* $f(x, y), g(x, y), h(x, y)$ *tre funzioni con un intorno di* $P_0 = (x_0, y_0)$*. Sia*

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

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

Questo perché la funzione centrale $g(x, y)$ è "intrappolata" tra la 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 θ (Funzione Radiale):** g deve dipendere esclusivamente dalla distanza ρ. Questo garantisce che il valore del limite non dipenda dalla direzione scelta per avvicinarsi al punto (x_0, y_0), assicurando l'uniformità della convergenza rispetto a θ;

- **Non negatività:** Deve risultare $g(\rho) \geq 0$, in quanto deve maggiorare un valore assoluto (che rappresenta una distanza geometrica tra il valore della funzione e il suo limite);

- **Natura infinitesima:** Deve essere verificato che $\lim_{\rho \to 0^+} g(\rho) = 0$.

La necessità di una funzione $g(\rho)$ non negativa e infinitesima risiede nell'applicazione del teorema del confronto.
Sapendo che il valore assoluto di una differenza è per definizione non negativo, possiamo impostare la seguente catena di 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 $\rho \to 0^+$, 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 $\lim_{\rho \to 0^+} g(\rho) = 0$.
3. Per il Teorema dei Carabinieri, si conclude che:

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

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

Studio della continuità

Sia $f : A \subseteq \mathbb{R}^2 \to \mathbb{R}$ 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 $\lim_{(x,y) \to (x_0, y_0)} f(x, y) = \ell$ 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 $\lim_{\rho \to 0^+} f(\rho, \theta) = 0$ e con le dovute analisi di maggiorazione si controlla che il limite dipenda uniformemente rispetto a θ verificando l'uguaglianza.

Calcolo Differenziale

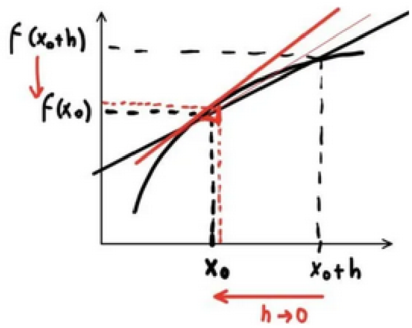
2.1 Derivata

2.1.1 Derivata parziale

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

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

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



Definizione 2.1.1 (Derivata parziale rispetto ad una variabile)**.** *La derivata parziale di una funzione $f : \mathbb{R}^n \to \mathbb{R}$ rispetto ad una variabile k è la derivata ordinaria di f considerata come funzione delle sola variabile k, mantenendo "fisse" le altre variabili.*

Il termine "parziale" indica che si sta considerando solo una parte del dominio, ovvero la variazione della funzione rispetto a una singola variabile, "congelando" tutte le altre. È una derivata parziale nel senso che cattura solo un aspetto del comportamento della funzione, non la sua variazione completa in tutte le direzioni.
La derivata parziale fornisce informazioni locali lungo una direzione specifica (quella dell'asse coordinata corrispondente), ma non descrive il comportamento globale della funzione in un intorno del punto.
Formalmente, sia un punto $P_0 = (x_0, y_0) \in A$ e sia $f : A \subseteq \mathbb{R}^2 \to \mathbb{R}$ con A aperto:

- La derivata rispetto a x (solo x incrementa)

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

- La derivata rispetto a y (solo y incrementa)

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

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 $y = y_0$:

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

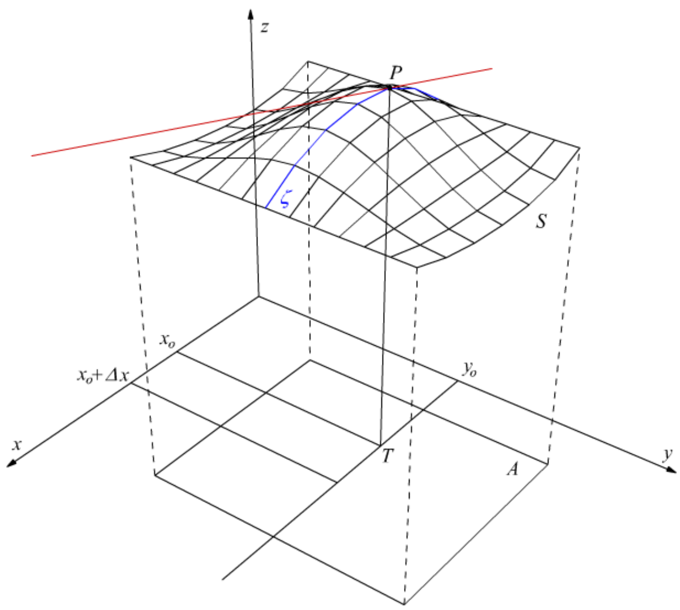
Definizione 2.1.2 (Derivata parziale rispetto a n variabili)**.** *Sia $f : A \subseteq \mathbb{R}^2 \to \mathbb{R}$ con A aperto e sia un punto $P_0 = (\underline{\boldsymbol{x}}) \in A$. La derivata parziale della f rispetto alla variabile x_i con $i = 1...n$ in $P_0 = (\underline{\boldsymbol{x}}) = (x_1...x_n)$ corrisponde a*

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

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

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

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



2.1.2 Gradiente

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

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

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

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

Affinché una funzione multivariata sia derivabile in un punto:

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

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

Definizione 2.1.3 (Gradiente)**.** *Il gradiente è un vettore denotato $\nabla f \in \mathbb{R}^n$ le cui componenti sono le derivate parziali della funzione f nel punto $\underline{\boldsymbol{x}}$:*

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

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:

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

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

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

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

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

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

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

  Quindi,

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

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

2.1.3 Derivata direzionale

Sia $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ e siano $\underline{\boldsymbol{x}} = \{x_1, ..., x_n\} \in A$ e e_i versore componenente della base canonica $e = (e_1, ..., e_n)$ di $\mathbb{R}^n$ (direzione nello spazio).
La derivata direzionale rispetto a f rispetto alla direzione di e_i nel punto $\underline{\boldsymbol{x}}$ se esiste allora esiste finito

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

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

La derivata parziale può essere rappresentata con un concetto più generale: la derivata direzionale. Dato un piano tangente, per il punto in cui si poggia il piano passano infinite rette.
Informalmente, la derivata direzionale è la pendenza di una di quelle rette in riferimento ad una direzione scelta partendo dal punto del dominio.

Definizione 2.1.4 (Derivata direzionale)**.** *Dato un versore $\underline{\boldsymbol{v}}$, ovvero un vettore di norma unitaria ($||v|| = 1$), in $\mathbb{R}^n$, sia $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ con A aperto e sia $\underline{\boldsymbol{x}} = (x_1...x_n) \in A$.*
La derivata direzionale di f rispetto alla direzione data da $\underline{\boldsymbol{v}}$ nel punto $\underline{\boldsymbol{x}}$ esiste se esiste ed è finito il limite

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

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 và normalizzato con la formula $v = \frac{v}{||v||}$.

2.2 Differenziabilità

Differenza tra derivabilità e differenziabilità

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

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

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

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

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

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

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

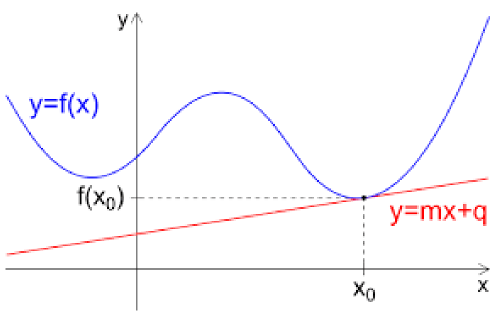
  Se dividiamo entrambi i membri per $x - x_0$ e prendiamo il limite per $x \to x_0$, otteniamo:

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

  Questo significa che:

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

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



- **Analisi 2 (Caso multivariato)**:

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

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

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

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

  dove

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

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

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

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

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

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

  tende sempre a zero, indipendentemente dalla direzione scelta.

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

  Esempio fondamentale (Derivabile ma non Differenziabile):

  Consideriamo la funzione

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

  Nel punto $(0, 0)$:

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

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

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

  Teorema 2.2.1 (Condizione sufficiente per la differenziabilità)**.** *Se tutte le derivate parziali di f esistono e sono <u>continue</u> in un intorno di $\underline{\boldsymbol{x}}_0$ (ovvero $f \in C^1$), allora f è differenziabile in $\underline{\boldsymbol{x}}_0$.*

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

2.2.1 Il concetto di funzione differenziabile e l'esistenza del piano tangente

Definizione 2.2.1 (Differenziabilità). *Una funzione $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ è differenziabile in $\underline{\boldsymbol{x}}_0 \in A$ se e solo se:*

- *f è derivabile in $\underline{\boldsymbol{x}}_0$, ovvero $\exists\ \nabla f(\underline{\boldsymbol{x}}_0)$;*
- *Vale il seguente limite:*

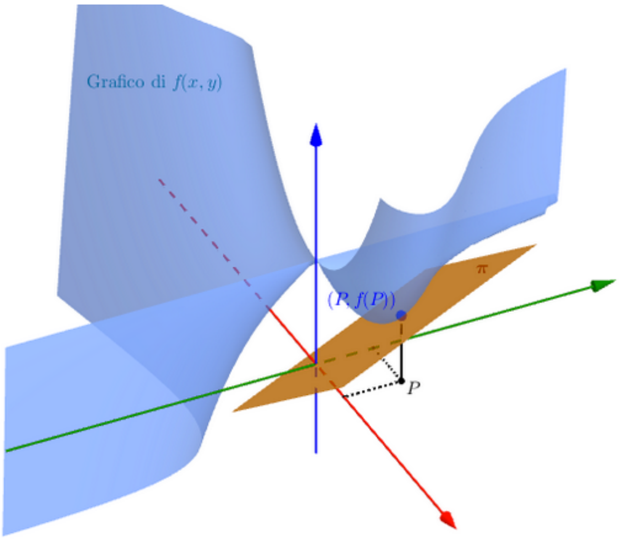
$$\lim_{\underline{\boldsymbol{h}} \to \underline{\boldsymbol{0}}} \frac{f(\underline{\boldsymbol{x}}_0 + \underline{\boldsymbol{h}}) - f(\underline{\boldsymbol{x}}_0) - \langle \nabla f(\underline{\boldsymbol{x}}_0) \cdot \underline{\boldsymbol{h}} \rangle}{\|\underline{\boldsymbol{h}}\|} = 0$$

Dove $\underline{\boldsymbol{h}} = (h_1, \ldots, h_n)$ è il vettore incremento e $\|\underline{\boldsymbol{h}}\| = \sqrt{\sum h_i^2}$ 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 $\|\underline{\boldsymbol{h}}\|$). In simboli:

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

Osservazione. Geometricamente, dire che una funzione $f : \mathbb{R}^2 \to \mathbb{R}$ è differenziabile in un punto $P_0 = (x_0, y_0)$ equivale a dire che il grafico della funzione ammetta un piano tangente non verticale in $P_0 = (\underline{\boldsymbol{x}}_0, f(\underline{\boldsymbol{x}}_0))$.
Non basta che il piano tocchi il punto (quello lo fanno infiniti piani). Significa che il piano è talmente "aderente" alla superficie curva che, se "zoomiamo" infinitamente sul punto P_0, la superficie curva del grafico diventa indistinguibile dal piano piatto. L'errore che commettiamo sostituendo la superficie con il piano tende a zero più velocemente della distanza dal punto.



Equazione del piano tangente

Sia $f : A \subseteq \mathbb{R}^2 \to \mathbb{R}$ differenziabile in $P_0 = (x_0, y_0)$. L'equazione del piano tangente al grafico in P_0 è:

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

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

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

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

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

Mentre l'approssimazione della funzione (formula di Taylor al primo ordine) è:

$$f(x) = \underbrace{f(x_0) + f'(x_0)(x - x_0)}_{\text{Parte Lineare (Retta)}} + \underbrace{o(x - x_0)}_{\text{Errore}} \quad \text{per } x \to x_0$$

Allo stesso modo, in $\mathbb{R}^2$, il piano tangente rappresenta la parte lineare dello sviluppo della funzione.

2.2.2 Continuità di una funzione differenziabile e il teorema del differenziale totale

Teorema 2.2.2 (Continuità di una funzione differenziabile)**.** *Sia $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ differenziabile in un punto $\underline{\boldsymbol{x}}_0 \in A$ allora è continua in $\underline{\boldsymbol{x}}_0$.*

DIMOSTRAZIONE

Per dimostrare la continuità in $\underline{\boldsymbol{x}}_0$, dobbiamo verificare che:

$$\lim_{\underline{\boldsymbol{h}} \to \underline{\boldsymbol{0}}} f(\underline{\boldsymbol{x}}_0 + \underline{\boldsymbol{h}}) = f(\underline{\boldsymbol{x}}_0)$$

Dalla definizione di differenziabilità, possiamo scrivere l'incremento della funzione come somma di una parte lineare e un infinitesimo di ordine superiore:

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

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

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

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 $\underline{\boldsymbol{h}}$ (non dipende dall'incremento), quindi:

$$\lim_{\underline{\boldsymbol{h}} \to \underline{\boldsymbol{0}}} f(\underline{\boldsymbol{x}}_0) = f(\underline{\boldsymbol{x}}_0)$$

- Il termine lineare (prodotto scalare) tende a 0. Possiamo dimostrarlo usando la disuguaglianza di Cauchy-Schwarz e il teorema del confronto. Notiamo che:

$$0 \leq |\nabla f(\underline{\boldsymbol{x}}_0) \cdot \underline{\boldsymbol{h}}| \leq \|\nabla f(\underline{\boldsymbol{x}}_0)\| \cdot \|\underline{\boldsymbol{h}}\|$$

  Poiché $\|\nabla f(\underline{\boldsymbol{x}}_0)\|$ è un numero finito e $\|\underline{\boldsymbol{h}}\| \to 0$, il prodotto tende a zero. Di conseguenza:

$$\lim_{\underline{\boldsymbol{h}} \to \underline{\boldsymbol{0}}} (\nabla f(\underline{\boldsymbol{x}}_0) \cdot \underline{\boldsymbol{h}}) = 0$$

- L'ultimo termine tende a zero per la definizione stessa di o-piccolo (errore infinitesimo):

$$\lim_{\underline{\boldsymbol{h}} \to \underline{\boldsymbol{0}}} o(\|\underline{\boldsymbol{h}}\|) = 0$$

Mettendo tutto insieme otteniamo:

$$\lim_{\underline{\boldsymbol{h}} \to \underline{\boldsymbol{0}}} f(\underline{\boldsymbol{x}}_0 + \underline{\boldsymbol{h}}) = f(\underline{\boldsymbol{x}}_0) + 0 + 0 = f(\underline{\boldsymbol{x}}_0)$$

Poiché il limite coincide con il valore della funzione nel punto, f è continua in $\underline{\boldsymbol{x}}_0$.

□

Il differenziale

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

- **L'incremento vero (Δf):** Rappresenta la variazione reale della funzione muovendosi sulla superficie curva del grafico.

$$\Delta f = f(\underline{\boldsymbol{x}}_0 + \underline{\boldsymbol{h}}) - f(\underline{\boldsymbol{x}}_0)$$

- **Il differenziale (df):** Rappresenta la variazione calcolata muovendosi sul piano tangente. È la parte lineare dell'incremento.

$$df = \nabla f(\underline{\boldsymbol{x}}_0) \cdot \underline{\boldsymbol{h}}$$

Mentre Δf misura il dislivello esatto sulla "montagna" (il grafico di f), il differenziale df misura il dislivello che avremmo se camminassimo sul piano tangente che approssima la montagna in quel punto.
Per spostamenti $\underline{\boldsymbol{h}}$ molto piccoli (in un intorno di $\underline{\boldsymbol{x}}_0$), 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 ($o(\|\underline{\boldsymbol{h}}\|)$).

Teorema 2.2.3 (Teorema del differenziale totale)**.** *Sia $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ con A aperto e sia f di classe C^1 su A, ovvero tutte le derivate parziali esistono e sono continue in un punto $\underline{\boldsymbol{x}} \in A$:*

$$\exists\ \nabla f(\underline{\boldsymbol{x}}) \Leftrightarrow \textit{esistono } f_{x_1}(\underline{\boldsymbol{x}}), f_{x_2}(\underline{\boldsymbol{x}}), \dots, f_{x_n}(\underline{\boldsymbol{x}})$$

Allora f è differenziabile in $\underline{\boldsymbol{x}}$.

DIMOSTRAZIONE su $n = 2$ (vale per n variabili)
Sia $f : A \subseteq \mathbb{R}^2 \to \mathbb{R}$ tale che $(x, y) \to f(x, y)$, si supppone 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 $(x, y), \forall \underline{\boldsymbol{h}} \in \mathbb{R}^2$ con $\underline{\boldsymbol{h}} = (h, k)$

$$\lim_{h,k \to \underline{\mathbf{0}}} \frac{f(x+h, y+k) - f(x, y) - \langle \nabla f(x, y) \cdot (h, k) \rangle}{\sqrt{h^2 + k^2}} = 0$$

Passo 1: Consideriamo l'incremento della funzione:

$$\Delta f = f(x+h, y+k) - f(x, y)$$

Aggiungiamo e sottraiamo il termine misto $f(x, y + k)$

$$= \underbrace{[f(x+h, y+k) - f(x, y+k)]}_{\text{Variazione solo in } x} + \underbrace{[f(x, y+k) - f(x, y)]}_{\text{Variazione solo in } y}$$

Perché nel primo membro la y in $y + k$ rimane costante, nel secondo membro invece è la x ad essere costante.

Passo 2: Applicazione del Teorema del valor medio (Lagrange)
Poiché f è derivabile, applichiamo il Teorema del valor medio separatamente alle due parentesi:

- Per la prima parentesi (variabile x), esiste un punto x_1 compreso tra x e $x + h$ tale che:

$$f(x+h, y+k) - f(x, y+k) = f_x(x_1, y+k) \cdot (x + h - x) = f_x(x_1, y+k) \cdot h \text{ con } h > 0$$

- Per la seconda parentesi (variabile y), esiste un punto y_1 compreso tra y e $y+k$ tale che:

$$f(x, y+k) - f(x, y) = f_y(x, y_1) \cdot (y+k-y) = f_y(x, y_1) \cdot k \text{ con } k > 0$$

Quindi l'incremento diventa:

$$\Delta f = f_x(x_1, y+k)h + f_y(x, y_1)k$$

Passo 3: Costruzione del limite di differenziabilità
Dobbiamo verificare che:

$$\lim_{(h,k)\to(0,0)} \frac{f(x+h, y+k) - f(x, y) - [f_x(x, y)h + f_y(x, y)k]}{\sqrt{h^2+k^2}} = 0$$

Sostituiamo Δf con l'espressione trovata con il teorema del valor medio. Il numeratore diventa:

$$\text{Num} = [f_x(x_1, y+k)h + f_y(x, y_1)k] - [f_x(x, y)h + f_y(x, y)k]$$

Raccogliamo h e k:

$$\text{Num} = [f_x(x_1, y+k) - f_x(x, y)]h + [f_y(x, y_1) - f_y(x, y)]k$$

Passo 4: Maggiorazione con il valore assoluto
Passiamo al valore assoluto e usiamo la disuguaglianza triangolare ($|a+b| \leq |a| + |b|$):

$$\left| \frac{\text{Num}}{\sqrt{h^2+k^2}} \right| \leq \frac{|f_x(x_1, y+k) - f_x(x, y)| \cdot |h|}{\sqrt{h^2+k^2}} + \frac{|f_y(x, y_1) - f_y(x, y)| \cdot |k|}{\sqrt{h^2+k^2}}$$

Osserviamo che i rapporti $\frac{|h|}{\sqrt{h^2+k^2}}$ e $\frac{|k|}{\sqrt{h^2+k^2}}$ sono sempre ≤ 1 (poiché il cateto è minore dell'ipotenusa). Quindi:

$$0 \leq \text{Limite} \leq |f_x(x_1, y+k) - f_x(x, y)| + |f_y(x, y_1) - f_y(x, y)|$$

Passo 5: Conclusione per continuità
Facciamo tendere $(h, k) \to (0, 0)$.

- Se $h \to 0$, allora $x_1 \to x$ (perché compreso tra x e $x+h$).
- Se $k \to 0$, allora $y_1 \to y$.

Poiché per ipotesi f è di classe C^1, le derivate parziali f_x e f_y sono <u>continue</u>. Pertanto:

$$\lim_{(h,k)\to(0,0)} f_x(x_1, y+k) = f_x(x, y)$$

$$\lim_{(h,k)\to(0,0)} |f_x(x_1, y+k) - f_x(x, y)| = 0$$

E

$$\lim_{(h,k)\to(0,0)} f_y(x, y_1) = f_y(x, y)$$

$$\lim_{(h,k)\to(0,0)} |f_y(x, y_1) - f_y(x, y)| = 0$$

Essendo dei valori assoluti tendono a zero quando $(h, k) \to (0, 0)$, quindi si può affermare che il limite è nullo, allora la funzione è differenziabile.

$\square$

Gerarchia di regolarità

Sia $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ con A aperto, definiamo $C^1(A)$ come l'insieme delle funzioni con derivate parziali continue su A.
Vale la seguente catena di implicazioni per funzioni di più variabili:

$$f \in C^1(A) \implies f \text{ differenziabile in } A \implies f \in C^0(A) \text{ continua in ogni punto di } A$$

$f \in C^1$ è una condizione *sufficiente* ma non necessaria per la differenziabilità. Esistono funzioni differenziabili che hanno derivate parziali discontinue (casi rari e patologici).

Relazione tra differenziabilità e derivate direzionali

Definizione 2.2.2 (La differenziabilità implica la continuità e l'esistenza delle derivate direzionali). *Sia $f : A \to \mathbb{R}$, con $A \subset \mathbb{R}^n$ aperto e $x_0 \in A$. Se f è differenziabile in $\underline{\boldsymbol{x}}_0$ con differenziale L, allora:*

1. *f è continua in $\underline{\boldsymbol{x}}_0$;*

2. *f ammette tutte le derivate direzionali in $\underline{\boldsymbol{x}}_0$ e*

$$D_{\vec{v}} f(\underline{\boldsymbol{x}}_0) = L(\vec{v}), \quad \forall \vec{v} \in \mathbb{R}^n.$$

DIMOSTRAZIONE
In particolare, si può dimostrare che il differenziale L, se esiste, è unico.
Sia $\underline{\boldsymbol{v}} \in \mathbb{R}^n$, per definizione, la derivata direzionale è:

$$D_{\underline{\boldsymbol{v}}} f(\underline{\boldsymbol{x}}_0) = \lim_{t \to 0} \frac{f(\underline{\boldsymbol{x}}_0 + t\underline{\boldsymbol{v}}) - f(\underline{\boldsymbol{x}}_0)}{t}$$

Usando la definizione di differenziabilità con incremento $\underline{\boldsymbol{h}} = t\underline{\boldsymbol{v}}$, si ottiene:

$$f(\underline{\boldsymbol{x}}_0 + t\underline{\boldsymbol{v}}) = f(\underline{\boldsymbol{x}}_0) + L(t\underline{\boldsymbol{v}}) + o(\|t\underline{\boldsymbol{v}}\|)$$

Per la linearità di L, sappiamo che $L(t\underline{\boldsymbol{v}}) = tL(\underline{\boldsymbol{v}})$. Sostituendo e portando $f(\underline{\boldsymbol{x}}_0)$ a sinistra:

$$f(\underline{\boldsymbol{x}}_0 + t\underline{\boldsymbol{v}}) - f(\underline{\boldsymbol{x}}_0) = tL(\underline{\boldsymbol{v}}) + o(|t|\|\underline{\boldsymbol{v}}\|)$$

Dividendo tutto per $t \neq 0$:

$$\frac{f(\underline{\boldsymbol{x}}_0 + t\underline{\boldsymbol{v}}) - f(\underline{\boldsymbol{x}}_0)}{t} = L(\underline{\boldsymbol{v}}) + \frac{o(|t|\|\underline{\boldsymbol{v}}\|)}{t}$$

Passando al limite per $t \to 0$, il termine con l'o-piccolo tende a 0. Segue che:

$$D_{\underline{\boldsymbol{v}}} f(\underline{\boldsymbol{x}}_0) = L(\underline{\boldsymbol{v}})$$

Poiché il limite (la derivata direzionale) è unico, anche l'applicazione lineare L deve essere unica.
□

Definizione 2.2.3 (La formula del gradiente). *Sia f differenziabile su A e sia un vettore $\vec{v}$ unitario, si può definire la derivata direzionale con la formula*

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

2.2.3 La regola della catena

Sia $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ una funzione differenziabile su un aperto A.
Sia $\gamma : I \subseteq \mathbb{R} \to A \subseteq \mathbb{R}^n$ una curva regolare (derivabile con derivata continua).
Allora la funzione composta

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

è derivabile in ogni punto $t \in I$ e la sua derivata la derivata equivale al prodotto scalare tra il gradiente di f e il vettore tangente alla curva γ:

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

nel caso $n = 2$ con $\gamma(t) = (x(t), y(t))$ si ha

$$F'(t) = f_x(x(t), y(t)) \cdot x'(t) + f_y(x(t), y(t)) \cdot y'(t)$$

DIMOSTRAZIONE
Passo 1: Consideriamo il rapporto incrementale di $F(t)$ in un punto $t \in I$ rispetto a un incremento $h \in \mathbb{R}$.
Poiché per ipotesi f è differenziabile nel punto $\underline{\boldsymbol{x}} = \gamma(t)$, per definizione di differenziabilità possiamo scrivere l'incremento di f come somma della parte lineare (gradiente) e dell'errore (o-piccolo). Allora:

$$f(\gamma(t+h)) - f(\gamma(t)) = \langle \nabla f(\gamma(t)), \gamma(t+h) - \gamma(t) \rangle + o(\|\gamma(t+h) - \gamma(t)\|)$$

Sostituiamo questa espressione nel rapporto incrementale:

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

Per la linearità del prodotto scalare, possiamo portare $1/h$ dentro il primo termine:

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

Passo 2: Passiamo ora al limite per $h \to 0$.
Poiché γ è derivabile in t, il rapporto incrementale vettoriale tende alla derivata $\gamma'(t)$:

$$\lim_{h \to 0} \frac{\gamma(t+h) - \gamma(t)}{h} = \gamma'(t)$$

Quindi il primo termine tende a:

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

Che è esattamente la tesi che cerchiamo.
Passo 3: Dobbiamo ora dimostrare che il resto tende a 0.

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

Supponendo sia non nullo, riscriviamo

$$\left| \frac{o(\|\gamma(t+h) - \gamma(t)\|)}{h} \right|$$

Moltiplichiamo e dividiamo per la norma dell'incremento $\|\gamma(t+h) - \gamma(t)\|$:

$$\lim_{h \to 0} \underbrace{\frac{o(\|\gamma(t+h) - \gamma(t)\|)}{\|\gamma(t+h) - \gamma(t)\|}}_{\text{(A)}} \cdot \underbrace{\frac{\|\gamma(t+h) - \gamma(t)\|}{|h|}}_{\text{(B)}}$$

Analizziamo i fattori:

- **Fattore (A):** Tende a 0 per definizione di o-piccolo. Infatti, per la continuità di γ, quando $h \to 0$ anche l'incremento $\|\gamma(t+h) - \gamma(t)\| \to 0$.
- **Fattore (B):**.

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

Tende alla norma della derivata $\|\gamma'(t)\|$, che è un numero reale finito (la derivata esiste per ipotesi)

Quindi abbiamo un termine che tende a 0 moltiplicato per una quantità limitata ($\|\gamma'(t)\|$). Il limite complessivo è 0.
Ricomponendo i pezzi, il limite del rapporto incrementale è:

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

□

Applicando la regola della catena possiamo rispondere alla domanda: in quale direzione si ha la massima crescita di una funzione differenziabile?
L'idea fondamentale è trasformare un problema in più variabili in un problema in una sola variabile, dove il concetto di pendenza è immediato.

Passo 1: Il punto e il movimento
Sia una funzione differenziabile $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ e sia $\underline{\boldsymbol{x}}_0 \in A$ un punto fissato.
Per parlare di pendenza dobbiamo chiarire come ci muoviamo a partire dal punto, allora consideriamo una curva regolare che passa per il punto $\underline{\boldsymbol{x}}_0$:

$$\gamma : [-\epsilon, \epsilon] \subset \mathbb{R} \to A \subseteq \mathbb{R}^n, \quad \gamma(0) = \underline{\boldsymbol{x}}_0$$

La curva rappresenta un sentiero nello spazio, definiamo pure il suo vettore velocità $\gamma'(0) = \underline{\boldsymbol{v}}$, che indica la direzione iniziale del movimento. Visto che ci interessa sola la direzione (e non la velocità) scriviamo $\underline{\boldsymbol{v}}$ come il vettore di direzione unitaria, $\|\underline{\boldsymbol{v}}\| = 1$.

Passo 2: Riduzione a una variabile
Definiamo la funzione composta (restrizione di f alla curva):

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

Questa funzione descrive il valore percorrendo la curva γ. Poiché F dipende da una sola variabile, la sua derivata $F'(0)$ rappresenta la pendenza della funzione lungo il sentiero scelto nel punto $t = 0$ (corrispondente a $\underline{\boldsymbol{x}}_0$).

Passo 3: La regola della catena
La differenziabilità di f ci garantisce che esiste una derivata unica in ogni direzione che approssima la funzione, ossia il gradiente. Allora possiamo applicare la regola della catena per calcolare

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

Se poniamo il vettore unitario $\underline{\boldsymbol{v}} = \gamma'(0)$, questa quantità coincide con la derivata direzionale di f in $\underline{\boldsymbol{x}}_0$ lungo la direzione $\underline{\boldsymbol{v}}$

$$\langle \nabla f(\underline{\boldsymbol{x}}_0), \underline{\boldsymbol{v}}\rangle = D_{\underline{\boldsymbol{v}}} f(\underline{\boldsymbol{x}}_0)$$

Dunque, la regola della catena ci dice che tutte le possibili pendenze di f sono governate dal prodotto scalare tra il gradiente e il vettore direzione.

Passo 4: Massimizzare la pendenza
Ricordando la definizione geometrica di prodotto scalare

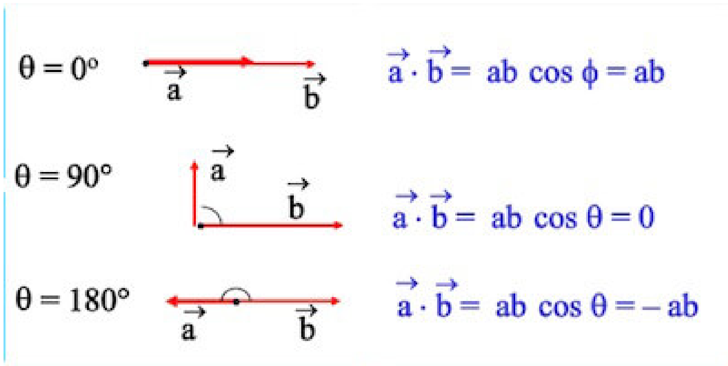
$$\underline{\boldsymbol{a}} \cdot \underline{\boldsymbol{b}} = \|\underline{\boldsymbol{a}}\|\|\underline{\boldsymbol{b}}\| \cos\theta$$

possiamo scrivere:

$$D_{\underline{v}} f = \|\nabla f(\underline{\boldsymbol{x}}_0)\| \cdot \underbrace{\|\underline{\boldsymbol{v}}\|}_{=1} \cdot \cos\theta = \|\nabla f(\underline{\boldsymbol{x}}_0)\| \cos\theta$$

dove θ è l'angolo tra il gradiente e il vettore unitario $\underline{\boldsymbol{v}}$, che rappresenta la direzione scelta.
Per massimizzare la pendenza $D_{\underline{v}} f$, dobbiamo massimizzare $\cos\theta$:

- **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,
$$D_{\underline{v}} f = -\|\nabla f\|$$
- **Crescita Nulla ($\theta = \pi/2$):** La funzione non cambia valore se ci muoviamo ortogonalmente al gradiente (lungo le curve di livello).



2.3 Derivate successive

Sia $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ funzione derivabile su A, ossia esistono e sono ben definite le derivate

$$\frac{\partial f}{\partial x_1}(\underline{\boldsymbol{x}}) \dots \frac{\partial f}{\partial x_n}(\underline{\boldsymbol{x}})$$

Si suppone che $\forall f_{x_i}(\underline{\boldsymbol{x}}) : A \to \mathbb{R}$ per qualche $i = 1 \dots n$ accade che la funzione $f_{x_i(\underline{\boldsymbol{x}})}$ sia derivabile ulteriormente, allora esiste la derivata successiva di f in $\underline{\boldsymbol{x}}$

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

La derivata seconda è, praticamente, la derivata parziale delle derivate prime di f e indicano la curvatura sulla superficie della funzione.

2.3.1 Matrice Hessiana e il teorema di Schwarz

Per raccogliere tutte le informazioni sulla curvatura della funzione, ci rifacciamo all'algebra lineare dove per ottenere uno scalare (un numero) moltiplicando dei vettori "due volte" (termini al quadrato tipo x^2 o misti xy), l 'oggetto centrale deve essere necessariamente una matrice quadrata.
Nel caso generale, le derivate seconde di una funzione f sono date dal variare degli indici i e j tra 1 e n. Quindi saranno $n \times n$ derivate seconde possibili organizzate in una Matrice Hessiana:

$$H_f(\underline{\boldsymbol{x}}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

In forma compatta si scrive anche

$$Hf(\underline{\boldsymbol{x}}) = \begin{bmatrix} f_{x_1x_1} & f_{x_1x_2} & \cdots & f_{x_1x_n} \\ f_{x_2x_1} & f_{x_2x_2} & \cdots & f_{x_2x_n} \\ \vdots & \vdots & \ddots & \vdots \\ f_{x_nx_1} & f_{x_nx_2} & \cdots & f_{x_nx_n} \end{bmatrix}$$

Se $i = j$ la derivata si dice pura e si scrive

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

Se, invece, $i \neq j$ la derivata si dice mista e si scrive

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

Matrice hessiana nel caso di $n = 2$

Sia $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ e supponiamo che f sia derivabile 2 volte nell'insieme A, ossia $\forall (x, y) \in A$ esistono tutte e 4 le derivate seconde:

$$f_{xx}(x, y), f_{xy}(x, y), f_{yx}(x, y), f_{yy}(x, y)$$

Queste derivate seconde sono organizzate in una matrice hessiana di dimensione 2×2

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

E' fondamentale affermare che, nel caso generale, <u>non</u> possiamo dire che la matrice hessiana sia simmetrica, solo dopo le opportune analisi dimostrando che appartiene alla classe C^2 potremo verificarlo.

Il teorema di Schwarz

Teorema 2.3.1 (Teorema di Schwarz)**.** *Sia $f : A \subseteq \mathbb{R}^2 \to \mathbb{R}$ 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 $D_{x_ix_j}(\underline{\boldsymbol{x}})$ $\forall i, j = 1 \ldots n$ e $\forall \underline{\boldsymbol{x}} \in A$. Se le derivate miste f_{xy} e f_{yx} esistono in un intorno di P_0 e sono continue in P_0, allora coincidono:*

$$f_{xy}(x_0, y_0) = f_{yx}(x_0, y_0)$$

Ovvero:

$$\frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) = \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0)$$

DIMOSTRAZIONE su $n = 2$
Sia $P = (x, y) \in A$ con $P \neq P_0$. Si definiscono le seguenti funzioni a partire da f:

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

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

$F(x)$ è la variazione della funzione sull'asse x quando y varia tra y e y_0, mentre $G(y)$ è la variazione della funzione sull'asse y quando x varia tra x e x_0.
Passo 1: Applicando il teorema del valor medio a $F(x)$ nell'intervallo di estremi x_0 e x, esiste un punto $x_1 \in (x_0, x)$ tale che:

$$F(x) - F(x_0) = F'(x_1)(x - x_0)$$

Poiché la derivata di F rispetto a x è data dalla differenza delle derivate parziali di f:

$$F'(x) = \frac{\partial f}{\partial x}(x, y) - \frac{\partial f}{\partial x}(x, y_0)$$

L'uguaglianza diventa:

$$F(x) - F(x_0) = \left(\frac{\partial f}{\partial x}(x_1, y) - \frac{\partial f}{\partial x}(x_1, y_0) \right) (x - x_0)$$

Applichiamo nuovamente il teorema del valor medio alla funzione $\frac{\partial f}{\partial x}$ (vista come funzione di y) nell'intervallo (y_0, y). Esiste un punto $y_1 \in (y_0, y)$ tale che:

$$\frac{\partial f}{\partial x}(x_1, y) - \frac{\partial f}{\partial x}(x_1, y_0) = \frac{\partial^2 f}{\partial y \partial x}(x_1, y_1)(y - y_0)$$

Sostituendo, otteniamo l'espressione finale per la differenza di F:

$$F(x) - F(x_0) = \frac{\partial^2 f}{\partial y \partial x}(x_1, y_1)(y - y_0)(x - x_0)$$

Analogamente, si applica lo stesso ragionamento alla funzione $G(y)$ nell'intervallo di estremi y_0 e y. Esiste un punto $y_2 \in (y_0, y)$ tale che:

$$G(y) - G(y_0) = G'(y_2)(y - y_0) = \left(\frac{\partial f}{\partial y}(x, y_2) - \frac{\partial f}{\partial y}(x_0, y_2) \right) (y - y_0)$$

Applicando il teorema del valor medio rispetto a x alla funzione $\frac{\partial f}{\partial y}$ nell'intervallo (x_0, x), esiste un punto $x_2 \in (x_0, x)$ tale che:

$$G(y) - G(y_0) = \frac{\partial^2 f}{\partial x \partial y}(x_2, y_2)(x - x_0)(y - y_0)$$

Passo 2: Osserviamo che le due differenze iniziali sono algebricamente identiche

$$F(x) - F(x_0) = [f(x, y) - f(x, y_0)] - [f(x_0, y) - f(x_0, y_0)]$$

$$G(y) - G(y_0) = [f(x, y) - f(x_0, y)] - [f(x, y_0) - f(x_0, y_0)]$$

Quindi $F(x) - F(x_0) = G(y) - G(y_0)$. Uguagliando le espressioni ottenute e dividendo per $(x - x_0)(y - y_0) \neq 0$:

$$\frac{\partial^2 f}{\partial y \partial x}(x_1, y_1) = \frac{\partial^2 f}{\partial x \partial y}(x_2, y_2)$$

Passo 3: Per l'ipotesi di continuità delle derivate miste in P_0, passando al limite per $(x, y) \to (x_0, y_0)$, i punti intermedi (x_1, y_1) e (x_2, y_2) tendono entrambi a (x_0, y_0):

$$\lim_{(x,y) \to (x_0, y_0)} \frac{\partial^2 f}{\partial y \partial x}(x_1, y_1) = \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0)$$

$$\lim_{(x,y)\to(x_0,y_0)} \frac{\partial^2 f}{\partial x \partial y}(x_2, y_2) = \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)$$

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.

$\square$

La sola esistenza delle derivate miste in un punto non garantisce l'uguaglianza.
È fondamentale l'ipotesi di continuità delle derivate seconde.
Esistono controesempi (es. la funzione di Peano-Schwarz $f(x,y) = \frac{xy(x^2-y^2)}{x^2+y^2}$) dove le derivate miste nell'origine esistono ma sono diverse ($f_{xy} \neq f_{yx}$) proprio perché non sono continue in $(0,0)$.

2.4 La formula di Taylor in funzioni multivariate

Per analizzarre il comportamento di una funzione multivariata $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ in un punto, possiamo utilizzare la formula di Taylor per approssimare la funzione con un polinomio.
Per calcolare la formula di Taylor, è necessario ridurre il problema a una variabile, utilizzando una parametrizzazione che ci consenta di studiare il comportamento della funzione lungo un segmento.
Sia $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ che si suppone sia di classe $C^2(A)$ e siano $\underline{\boldsymbol{x}}, \underline{\boldsymbol{x}} + \underline{\boldsymbol{h}} \in A \subseteq \mathbb{R}^n$ con $\underline{\boldsymbol{h}} \neq 0$.
Possiamo dire che parametrizzare il segmento di estremi $\underline{\boldsymbol{x}}$ e $\underline{\boldsymbol{x}} + \underline{\boldsymbol{h}}$ contenuto in A:

$$[\underline{\boldsymbol{x}}, \underline{\boldsymbol{x}} + \underline{\boldsymbol{h}}] = \{\underline{\boldsymbol{x}}(t) \in \mathbb{R}^n : \underline{\boldsymbol{x}}(t) = \underline{\boldsymbol{x}} + t\underline{\boldsymbol{h}},\ t \in [0,1]\} \subset A.$$

Si osserva che la parametrizzazione soddisfa le condizioni agli estremi:

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

Ora possiamo scrivere la funzione ausiliaria (restrizione di f al segmento), sia $F : [0,1] \to \mathbb{R}$:

$$F(t) = f(\underline{\boldsymbol{x}} + t\underline{\boldsymbol{h}}), \quad \forall t \in [0,1].$$

Il nostro obiettivo, adesso, è quello di dimostrare che dalla formula di Taylor centrata in t possiamo ricavare informazioni su $f(\underline{\boldsymbol{x}})$.

Se $f \in C^1(A)$, allora $F \in C^1([0,1])$.
La derivata prima $F'(t)$ conla regola della catena:

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

Poiché abbiamo supposto che $f \in C^2(A)$, applicando nuovamente la regola della catena a ciascun termine di $\frac{\partial f}{\partial x_i}$ otteniamo:

$$F''(t) = \sum_{i,j=1}^{n} \frac{\partial f}{\partial x_j}\Big(\frac{\partial f}{\partial x_i}(\underline{\boldsymbol{x}} + t\underline{\boldsymbol{h}}) \cdot h_i\Big) \cdot h_j = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i, \partial x_j}(\underline{\boldsymbol{x}} + t\underline{\boldsymbol{h}}) \cdot h_i \cdot h_j$$

<u>Formalmente</u>, è un polinomio omogeneo di secondo grado con le componenti del vettore $\underline{\boldsymbol{h}}$ che appaiono moltiplicate fra di loro. I coefficenti di questo polinomio sono proprio le derivate seconde della funzione in $(\underline{\boldsymbol{x}}, \underline{\boldsymbol{x}} + t\underline{\boldsymbol{h}})$:

- $F(0) = f(\underline{\boldsymbol{x}}(0)) = f(\underline{\boldsymbol{x}}) = \underline{\boldsymbol{x}}$

- $F(1) = f(\underline{\boldsymbol{x}}(1)) = f(\underline{\boldsymbol{x}} + \underline{\boldsymbol{h}}) = \underline{\boldsymbol{x}} + \underline{\boldsymbol{h}}$

Allora applicando la formula di Taylor alla $F(t)$ centrata nel punto $t = 0$ ottengo

$$F(t) = P_n(t, 0) + R_n(t, 0)$$

Generalizzando, se la funzione $f \in C^k(A)$ allora trasferisce la sua regolarità su una composta F. Possiamo affermare che $\exists\ \theta \in \mathbb{R}$ che sta tra $[0, 1]$ con resto tale che

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

<u>Informalmente</u>, invece di studiare il comportamento della funzione f nell'intero dominio multidimensionale $A \subseteq \mathbb{R}$, stiamo "affettando" il grafico della funzione lungo una retta. Definendo $F(t) = f(\underline{\boldsymbol{x}} + t\underline{\boldsymbol{h}})$, stiamo restringendo il nostro campo d'osservazione a ciò che accade esclusivamente lungo il segmento $(x, x + h)$.
I punti chiave di questa osservazione sono:

- **La Funzione come Binario**

  La variabile scalare t agisce come un parametro temporale che ci fa scorrere lungo un binario rettilineo. Questo ci permette di ricondurre lo studio di una funzione di n variabili allo studio di una funzione di una sola variabile reale ($F : [0, 1] \to \mathbb{R}$);

- **La Derivata Direzionale**

  Il risultato $F'(t) = \langle \nabla f(\underline{\boldsymbol{x}} + t\underline{\boldsymbol{h}}), \underline{\boldsymbol{h}} \rangle$ non è altro che la derivata direzionale di f lungo il vettore $\underline{\boldsymbol{h}}$ (a meno di un fattore di normalizzazione). Ci dice quanto velocemente cambia il valore della funzione mentre ci muoviamo in quella specifica direzione $\underline{\boldsymbol{h}}$, partendo dal punto $\underline{\boldsymbol{x}} + t\underline{\boldsymbol{h}}$.

- **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.

2.4.1 La formula di Taylor secondo il resto di Lagrange e di Peano

Teorema 2.4.1 (Formula di Taylor con resto di Lagrange)**.** *Sia $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ con A un aperto e $f \in C^k(A)$.*
Sia il segmento di estremi $\underline{\boldsymbol{x}}$ e $\underline{\boldsymbol{x}} + \underline{\boldsymbol{h}}$ contenuto in A, allora si può trovare un punto θ reale $\in [0, 1]$ tale che il polinomio di Taylor centrato in $\underline{\boldsymbol{x}}$ con resto secondo Lagrange di ordine 1 è:

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

che equivale alla formula di Taylor $T_n \Rightarrow F(1) = F(0) + F'(0)$.
Nel caso in cui l'ordine sia l'ordine sia 2:

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

che in forma compatta viene $f(\underline{\boldsymbol{x}}) + \langle \nabla f(\underline{\boldsymbol{x}}), \underline{\boldsymbol{h}} \rangle + \frac{1}{2}\langle Hf(\underline{\boldsymbol{x}} + \theta\underline{\boldsymbol{h}}), \underline{\boldsymbol{h}} \rangle$.

DIMOSTRAZIONE del caso k = 2
La matrice Hessiana in ogni punto di A è una matrice $n \times n$:

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

In termini matriciali:

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

Esplicitamente la matrice Hessiana è:

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

□

Diversa dalla formula con il resto di Lagrange è la formula con il resto di Peano perché si considera il gradiente e la matrice hessiana centrati in $\underline{\boldsymbol{x}}$ e non in un punto intermedio ignoto θ.

Teorema 2.4.2 (Formula di Taylor con resto di Peano)**.** *Sia* $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ *e* $f \in C^2(A)$. *Allora:*

$$f(\underline{\boldsymbol{x}} + \underline{\boldsymbol{h}}) = f(\underline{\boldsymbol{x}}) + \langle \nabla f(\underline{\boldsymbol{x}}), \underline{\boldsymbol{h}} \rangle + \frac{1}{2} \langle Hf(\underline{\boldsymbol{x}})\underline{\boldsymbol{h}}, \underline{\boldsymbol{h}} \rangle + o(\|h\|^2) \quad \textit{per } \underline{\boldsymbol{h}} \to \underline{\boldsymbol{0}}$$

DIMOSTRAZIONE del caso n = 2
Sia $f : A \subseteq \mathbb{R}^2 \to \mathbb{R}$. Consideriamo i 2 punti $(\underline{\boldsymbol{x}}_0, \underline{\boldsymbol{y}}_0)$ e l'incremento $\underline{\boldsymbol{h}}$.

$$\underline{\boldsymbol{x}} = (x_0, y_0), \quad \underline{\boldsymbol{h}} = (h, k) \in \mathbb{R}^2$$

Quindi il punto incrementato è:

$$\underline{\boldsymbol{x}} + \underline{\boldsymbol{h}} = (x_0 + h, y_0 + k)$$

La formula si scrive:

$$f(x_0 + h, y_0 + k) = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)h + \frac{\partial f}{\partial y}(x_0, y_0)k + \frac{1}{2} \langle Hf(x_0, y_0)\underline{\boldsymbol{h}}, \underline{\boldsymbol{h}} \rangle$$

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 $\underline{\boldsymbol{h}} = \binom{h}{k}$. Prima moltiplicazione (Matrice × 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 $\underline{\boldsymbol{h}}^T$ (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:

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

Tutto sommato per l'errore:

$$+o(\sqrt{h^2+k^2}^2) = o(h^2+k^2) \quad \text{con } h,k \to 0$$

Riscrivendo in modo opportuno ponendo $x = x_0 + h \implies h = x - x_0$ e $y = y_0 + k \implies k = y - y_0$:

$$P_2(x,y) = f(x_0,y_0) + \frac{\partial f}{\partial x}(x-x_0) + \frac{\partial f}{\partial y}(y-y_0) + \frac{1}{2}\left[f_{xx}(x-x_0)^2 + 2f_{xy}(x-x_0)(y-y_0) + f_{yy}(y-y_0)^2\right]$$

□

Questo è l'unico polinomio in x,y di grado ≤ 2 per cui si ha che la differenza tra $f(x,y)$ e P_2 va a zero più rapidamente della distanza al quadrato per i 2 punti $o\|((x-x_0)^2 + (y-y_0)^2)\|$.

$$\lim_{(x,y)\to(x_0,y_0)} \frac{f(x,y) - P_2(x,y)}{\|(x,y)-(x_0,y_0)\|^2} = 0$$

Nel caso n variabili (caso generale), il termine $\langle Hf(x_0)h, h\rangle$ "esce fuori" un'espressione omogenea di 2° grado in $h_1, \ldots, h_n$, dove n è il numero di variabili, e si indicherà come $q(h)$, i cui coefficienti sono le opportune derivate seconde.

2.5 Studio di massimi e minimi per le funzioni a più variabili

Sia $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ una funzione definita su un aperto A e sia $\underline{\boldsymbol{x}}_0 \in A$ un punto di A.

Definizione 2.5.1 (Massimo e minimo locale). *Se $\exists\ B(\underline{\boldsymbol{x}}_0, r)$ con $r > 0$ (intorno di $\underline{\boldsymbol{x}}_0$) per cui $\forall x \in A \cap B(\underline{\boldsymbol{x}}_0, r)$ $f(\underline{\boldsymbol{x}}_0) \leq f(\underline{\boldsymbol{x}})$ allora $\underline{\boldsymbol{x}}_0$ è un punto di minimo locale di f, mentre se $f(\underline{\boldsymbol{x}}_0) \geq f(\underline{\boldsymbol{x}})$ allora $\underline{\boldsymbol{x}}_0$ è un punto di massimo locale di f.*

Si dice massimo (minimo) forte se la disuguaglianza è stretta, per $\underline{\boldsymbol{x}} \neq \underline{\boldsymbol{x}}_0$.
I punti di massimo e minimo relativi si chiamano anche punti di estremo locale.

2.6 Ottimizzazione libera

L'ottimizzazione libera si occupa di trovare i massimi e minimi di una funzione senza alcuna restrizione sul dominio. In questo caso, si cerca di identificare i punti in cui la funzione raggiunge i suoi valori estremi all'interno dell'intero spazio definito dalla funzione stessa.

Teorema 2.6.1 (Teorema di Fermat). *Sia f definita su un dominio $D = A \cup \partial A$ con A aperto, allora $f : D \subset \mathbb{R}^n \to \mathbb{R}$ e sia $\underline{\boldsymbol{x}}_0 \in D$ o sia $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ e $\underline{\boldsymbol{x}}_0 \in A$ un punto di estremo locale per f.*
Se f è differenziabile in tale punto $\underline{\boldsymbol{x}}_0$, il gradiente di f in $\underline{\boldsymbol{x}}_0$ è nullo.

Questa è una condizione necessaria ma non sufficiente per l'esistenza di massimi e minimi locali perché esistono punti, detti critici o stazionari, che hanno gradiente nullo ma non sono né massimi né minimi (es. punti di sella o colle).
DIMOSTRAZIONE
Supponiamo che $\underline{\boldsymbol{x}}_0$ sia un punto di massimo locale (la dimostrazione per il minimo è analoga).
Passo 1: Definizione della restrizione
Consideriamo la restrizione della funzione f lungo una direzione arbitraria governata dal versore $\underline{\boldsymbol{v}}$.

Definiamo la funzione scalare di una sola variabile $F(t)$:

$$F(t) = f(\underline{\boldsymbol{x}}_0 + t \cdot \underline{\boldsymbol{v}})$$

dove $\underline{\boldsymbol{v}}$ è un versore e t varia in un intorno $(-\delta, \delta)$.
Passo 2: Osservazione sul massimo.
Poiché $\underline{\boldsymbol{x}}_0$ è un punto di massimo locale per f, allora $f(\underline{\boldsymbol{x}}_0) \geq f(\underline{\boldsymbol{x}})$ per ogni $\underline{\boldsymbol{x}}$ in un intorno sferico.
Di conseguenza, in un intorno piccolo di t restringendoci alla retta:

$$F(0) = f(\underline{\boldsymbol{x}}_0) \geq f(\underline{\boldsymbol{x}}_0 + t \cdot \underline{\boldsymbol{v}}) = F(t)$$

Quindi, $t = 0$ è un punto di massimo locale per la funzione $F(t)$.
Passo 3: Applicazione del Teorema di Fermat (Analisi I).
Applicando il classico Teorema di Fermat per funzioni di una variabile, sappiamo che la derivata nel punto di massimo interno deve annullarsi:

$$F'(0) = 0$$

Passo 4: Legame con le derivate parziali.
Calcoliamo $F'(t)$ usando la regola della catena:

$$F'(t) = \langle \nabla f(\underline{\boldsymbol{x}}_0 + t \cdot \underline{\boldsymbol{v}}), \underline{\boldsymbol{v}} \rangle$$

Valutando in $t = 0$:

$$F'(0) = \langle \nabla f(\underline{\boldsymbol{x}}_0), \underline{\boldsymbol{v}} \rangle \iff \nabla f(\underline{\boldsymbol{x}}_0) = 0 \ \forall \ \underline{\boldsymbol{v}} \in \mathbb{R}^n$$

□

2.6.1 Test dell'hessiana

Per le funzioni ad una variabile, supposto che una funzione $f \in C^2$ e individuato un punto critico x_0 ($f'(x_0) = 0$), la classificazione del punto si ottiene dallo studio della derivata seconda, 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 $f : A \subseteq \mathbb{R}^2 \to \mathbb{R}$ 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 $\nabla f(P_0) = \underline{\mathbf{0}}$ 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.

2.6.2 Forme Quadratiche

In un punto critico la parte lineare dello sviluppo di Taylor si annulla (il piano tangente è piatto). Per studiare il comportamento locale della funzione è quindi necessario analizzare il termine successivo, quello di secondo ordine. Quindi, per capire l'origine del test precedente, analizziamo la formula di Taylor al secondo ordine.
Il segno della variazione locale della funzione dipende dal termine di secondo grado:

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

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

Definizione 2.6.1 (Forma Quadratica)**.** *Una forma quadratica in $\underline{\boldsymbol{h}} \in \mathbb{R}^n$ è un polinomio omogeneo di secondo grado nelle variabili $h_1,\dots,h_n$. Ad essa $q(\underline{\boldsymbol{h}})$ è associata una matrice simmetrica A tale che:*

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

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

Per esempio, dato il polinomio omogeneo:

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

La matrice simmetrica associata A è:

$$A=\begin{bmatrix}3 & 5\\ 5 & 27\end{bmatrix}$$

Notare come il coefficiente 10 di h_1h_2 sia stato diviso equamente tra a_{12} e a_{21}.
Per determinare la natura del punto critico, dobbiamo studiare il segno di $q(\underline{\boldsymbol{h}})$ per $\underline{\boldsymbol{h}} \neq \underline{\boldsymbol{0}}$.
Sia A la matrice simmetrica reale, essa ammette n autovalori reali $\lambda_1,\dots,\lambda_n$.

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

Nel caso di 2 variabili, si può dire che gli autovalori λ_1,λ_2 sono le radici del polinomio caratteristico

$$\lambda^2-\mathrm{tr}(A)\lambda+\det(A)=0$$

E valgono le relazioni

$$\det(A)=\lambda_1\cdot\lambda_2 \quad ; \quad \mathrm{tr}(A)=\lambda_1+\lambda_2$$

Quindi la matrice simmetrica A di dimensione 2×2 associata alla forma quadratica $q(h,k)$ è:

$$A=\begin{bmatrix}a_{11}-\lambda & a_{12}\\ a_{12} & a_{22}-\lambda\end{bmatrix} = \lambda^2-\underbrace{(a_{11}+a_{22})}_{\mathrm{tr}(A)}\lambda+\underbrace{a_{11}\cdot a_{22}-a_{12}^2}_{\det(A)}=0$$

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.

Quindi, il problema geometrico della classificazione dei punti critici viene ricondotto a un problema algebrico: lo studio del segno della forma quadratica associata al termine di secondo ordine dello sviluppo di Taylor.
Possiamo riassumere il ragionamento in tre passaggi chiave:

1. **Fermat non basta**

   Quando siamo in un punto critico ($\nabla f = 0$), il piano tangente è orizzontale. La parte lineare della funzione si annulla. Per capire se siamo su una cima, in una buca o su una sella, dobbiamo guardare il "termine successivo" dello sviluppo di Taylor: il termine di secondo grado (quadratico).

2. **Gli autovalori sono le "curvature vere"**

   La matrice Hessiana descrive la curvatura, ma spesso contiene "termini misti" (f_{xy}) che creano confusione perché dipendono dal sistema di coordinate scelto (x, y). Gli autovalori (λ) rappresentano le curvature lungo le direzioni principali, ovvero quelle direzioni speciali in cui la superficie curva senza torsioni.

   - Un λ positivo indica una curvatura verso l'alto (convessità).
   - Un λ negativo indica una curvatura verso il basso (concavità).

3. **Hessiana vs Autovalori**

   A prima vista può sembrare strano: gli autovalori sono così importanti per classificare i punti critici, eppure quando scriviamo l'Hessiana

   $$H = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{bmatrix}$$

   non vediamo λ_1 e λ_2 da nessuna parte.

   La chiave è capire che l'Hessiana "vede" la geometria con gli occhiali delle coordinate (x,y) scelte. Le derivate seconde f_{xx}, f_{xy}, f_{yy} descrivono come la funzione si curva lungo gli assi x e y e nelle direzioni miste. Ma queste direzioni sono arbitrarie: dipendono da come abbiamo orientato il nostro sistema di riferimento.

   Gli autovalori, invece, sono invarianti geometrici: non dipendono dal sistema di coordinate. Essi rappresentano le curvature lungo le direzioni principali del 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 $\implies$ minimo;
   - $\lambda_1 < 0, \lambda_2 < 0$ la parabola che scende in entrambe le direzioni $\implies$ massimo;
   - $\lambda_1 \cdot \lambda_2 < 0$ la parabola sale in una direzione e scende nell'altra $\implies$ sella

2.6.3 Caso dubbio: il test dell'hessiana non funziona

Quando il determinante è nullo, la matrice Hessiana è "degenere" e non ci dà informazioni sufficienti sulla curvatura (la funzione è "troppo piatta" per l'approssimazione quadratica).
Infatti, il metodo dell'Hessiana è inconcludente se abbiamo almeno un autovalore uguale a zero: non possiamo capire in modo certo se quel punto è di massimo, di minimo o di sella, e dobbiamo quindi ricorrere a un metodo alternativo.
Nell'analisi unidimensionale possiamo immaginare di voler classificare un punto critico x_1 di una funzione $f(x)$ sapendo solo che $f(x_1) = c$. Guardando il grafico sapremmo riconoscere immediatamente se x_1 è un massimo o un minimo.
Senza accesso al grafico, possiamo traslare verticalmente la funzione in modo che l'estremo si trovi esattamente sull'asse delle x, definiamo quindi

$$\bar{f}(x) = f(x) - c$$

Questa nuova funzione è identica a quella originale, ma "abbassata" (o "innalzata", se $c < 0$) esattamente del valore c. In particolare:

$$\bar{f}(x_1) = f(x_1) - c = 0$$

Ora studiamo il segno di $\bar{f}$ in un intorno di x_1:

- Se $\bar{f}(x) \leq 0$ in un intorno di x_1, significa che dopo la traslazione tutti i valori della funzione vicino a x_1 sono negativi o nulli.

  Quindi $f(x) \leq c$ in quell'intorno $\implies$ x_1 è un punto di massimo.
- Se $\bar{f}(x) \geq 0$ in un intorno di x_1, significa che $f(x) \geq c$ in quell'intorno $\implies$ x_1 è un punto di minimo.
- Se $\bar{f}$ cambia segno, allora x_1 non è né un massimo né un minimo (caso che in una variabile è raro, ma in più variabili corrisponde alla sella).

Questo stesso ragionamento si applica alle funzioni di due variabili quando l'Hessiana è degenere.
In questo caso, si deve tornare alla definizione di estremo locale, studiando direttamente il segno della variazione della funzione in un intorno del punto critico $P_0 = (x_0, y_0)$
Si definisce l'incremento:

$$\Delta f(x, y) = f(x, y) - f(x_0, y_0)$$

Questa è esattamente l'analogo bidimensionale della funzione $\bar{f}$ vista prima: stiamo "traslando" la superficie in modo che il punto critico si trovi a quota zero, e poi studiamo se la superficie nei dintorni è tutta sopra, tutta sotto, o a cavallo del piano $z = 0$.
Si studia quindi il segno di Δf in un intorno sufficientemente piccolo di P_0:

1. **Minimo Locale:** Se esiste un intorno in cui $\Delta f(x, y) \geq 0$ per ogni punto (ovvero la funzione assume sempre valori maggiori o uguali a $f(P_0)$);
2. **Massimo Locale:** Se esiste un intorno in cui $\Delta f(x, y) \leq 0$ per ogni punto (la superficie "traslata" è tutta sotto il piano $z = 0$);
3. **Punto di Sella:** Se in qualsiasi intorno di P_0, per quanto piccolo, esistono sia punti in cui $\Delta f > 0$ sia punti in cui $\Delta f < 0$. Graficamente, studiando il segno sul piano xy, vedremo delle regioni positive e negative che si alternano e convergono tutte nel punto critico.

2.7 Ottimizzazione vincolata

A differenza dell'ottimizzazione libera, dove cerchiamo estremi su un aperto $A \subseteq \mathbb{R}^n$, qui cerchiamo i massimi e minimi di una funzione $f(x, y)$ limitatamente ai punti che appartengono a un insieme V (detto *vincolo*).

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

In genere V è una curva di livello (es. una circonferenza, una retta) definita implicitamente da un'equazione $g(x, y) = 0$ (o k).
Quindi risulta possibile esplicitare una funzione rispetto ad un'altra (teorema della funzione implicita):

$$y = y(x) \text{ allora il vincolo diventa } g(x, y(x)), \text{ una funzione ad una variabile}$$

Se si studia la natura dei punti critci quando l'insieme di definizione è compatto, si può usare il Teorema di Weierstrass.

Teorema 2.7.1 (Teorema di Weierstrass)**.** *Sia $f : K \subseteq \mathbb{R}^n \to \mathbb{R}$ 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:

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

2.7.1 Parametrizzazione

Se l'equazione del vincolo $g(x, y) = 0$ è semplice, possiamo parametrizzare la curva $(x(t), y(t))$ trasformando il vincolo in un sostegno ad una curva $V = (\gamma[a, b])$.

$$\begin{cases} x = x(t) \\ y = y(t) \end{cases} \qquad t \in [0, h]$$

Questo metodo è comodo per vincoli lineari (es. $3x + 5y = 7$) ma difficile per vincoli complessi.

2.7.2 Moltiplicatori di Lagrange

Siano due funzioni $f, g : A \subseteq \mathbb{R}^n \to \mathbb{R}$ e siano entrambe appartenenti alla classe $C^1(A)$. e sia l'insieme $V = \{(x, y) \in A : g(x, y) = k\}$.
Supponenendo che $\underline{\boldsymbol{x_0}} \in V$ sia un punto di estremo locale per f ristretto all'insieme V e che il gradiente $\nabla g(\underline{\boldsymbol{x_0}}) \neq \underline{\mathbf{0}}$ (regolare rispetto in quel punto), allora esiste un numero reale $\lambda \in \mathbb{R}$ (detto moltiplicatore di lagrange) tale che:

$$\nabla f(x, y) = \lambda \nabla g(x, y)$$

Geometricamente, significa che nel punto di estremo le curve di livello di f sono tangenti al vincolo V (i gradienti sono paralleli).
Per trovare i candidati punti di estremo vincolato definiamo una funzione detta <u>Lagrangiana</u>:

$$\mathcal{L}(x, y, \lambda) = f(x, y) - \lambda[g(x, y) - k]$$

che ha la stessa regolarità della funzione di cui è composta.
I punti critici vincolati si trovano cercando i punti stazionari di $\mathcal{L}$ (dove il gradiente "completo" si annulla). Si risolve il sistema:

$$\nabla \mathcal{L}(x, y, \lambda) = \underline{\mathbf{0}} \Leftrightarrow \begin{cases} \frac{\partial f}{\partial x} - \lambda g_x = 0 \\ \frac{\partial f}{\partial y} - \lambda g_y = 0 \\ g(x, y) - k = 0 \quad \text{(la derivata del vincolo rispetto a } \lambda) \end{cases}$$

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

2.7.3 Esempio di ottimizzazione vincolata con Sostituzione e Lagrange.

Trovare max e min di $f(x, y) = (x-1)^2 - y^2$ *soggetta al vincolo* $x^2 + y^2 = 1$.
Il vincolo è la circonferenza unitaria (compatto $\to$ Weierstrass garantisce l'esistenza).
Metodo 1 (Sostituzione): Dal vincolo: $y^2 = 1 - x^2$ (con $x \in [-1, 1]$). Sostituiamo in f:

$$h(x) = (x-1)^2 - (1-x^2) = x^2 - 2x + 1 - 1 + x^2 = 2x^2 - 2x$$

Derivata: $h'(x) = 4x - 2 \implies x = 1/2$. Confrontiamo i valori in $x = 1/2$ e agli estremi $x = \pm 1$:

- $x = 1 \implies y = 0 \implies f(1, 0) = 0$
- $x = -1 \implies y = 0 \implies f(-1, 0) = 4$ (Max)
- $x = 1/2 \implies y = \pm\sqrt{3}/2 \implies f(1/2, \pm\sqrt{3}/2) = -1/2$ (Min)

Metodo 2 (Lagrange): Costruiamo $\mathcal{L}(x, y, \lambda) = (x-1)^2 - y^2 - \lambda(x^2 + y^2 - 1)$. Sistema:

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

Dalla seconda equazione: $y = 0$ o $\lambda = -1$.

- Se $y = 0$: dal vincolo $x = \pm 1$. Ottengo i punti $(1, 0)$ e $(-1, 0)$.
- Se $\lambda = -1$: dalla prima equazione $2(x-1) = -2x \implies 2x - 2 = -2x \implies 4x = 2 \implies x = 1/2$. Dal vincolo $y^2 = 1 - (1/4) = 3/4 \implies y = \pm\sqrt{3}/2$.

Si ritrovano gli stessi 4 punti candidati.

2.8 Integrazione

Mentre l'integrale dell'analisi unidimensionale $\int_a^b f(x)dx$ calcola l'area tra la curva e l'asse x, l'integrale doppio $\iint_R f(x,y)\,dx\,dy$ calcola il volume compreso tra la superficie definita da $z = f(x,y)$ e il piano xy sopra un dominio R nel piano.

Sia una funzione $f : R \to \mathbb{R}$ 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:

$$D_1 = \{a = x_0, x_1, \dots, x_n = b\} \qquad D_2 = \{c = y_0, y_1, \dots, y_m = d\}$$

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 $m \le f(x,y) \le M$ 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

$$\sum_{i,j=1}^{n,m} m_{i,j} \cdot \underbrace{(x_i - x_{i-1}) \cdot (y_j - y_{j-1})}_{\text{Area}(R_{i,j})}$$

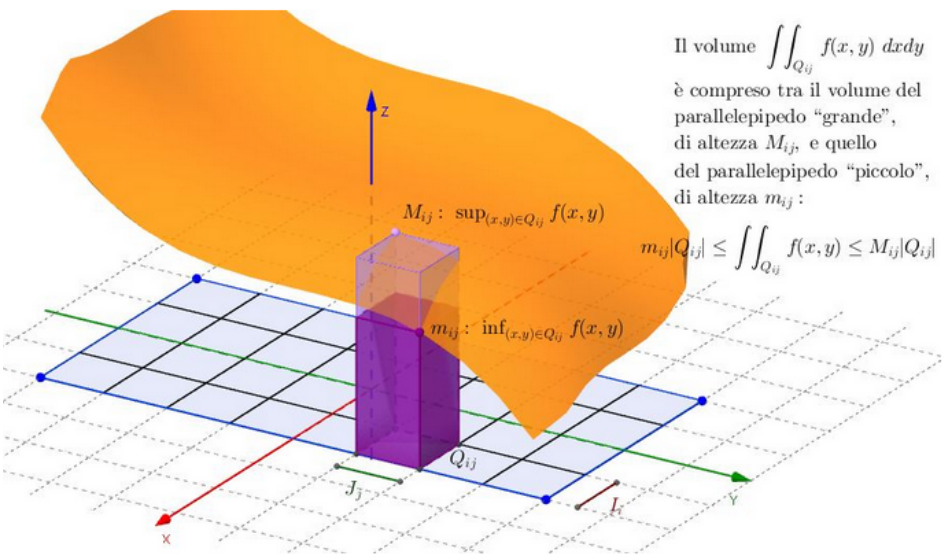
La somma di tutti questi volumi viene definita come la somma inferiore: $s = (f, D)$.
Stesso ragionamento vale per la somma superiore $S(f,D)$, dove si considera $M_{i,j} = sup f(x,y)$ come l'estremo superiore (il valore massimo) assunto dalla $f(x,y)$.

$$\sum_{i,j=1}^{n,m} M_{i,j} \cdot \underbrace{(x_i - x_{i-1}) \cdot (y_j - y_{j-1})}_{\text{Area}(R_{i,j})}$$

Queste due somme sono legate dalla relazione proveniente dal teorema della media integrale

$$m(b-a)(d-c) \le s \le S \le M(b-a)(d-c) \quad \forall \text{ suddivisione di } R$$

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 unidemnsionale si dimostra che vale la relazione per cui $\sup_D s(f,D) \le \inf_D S(f,D)$, che nel caso di una uguaglianza ci permette di definire una funzione integrabile.



2.8.1 Integrale doppio e concetto di continuità

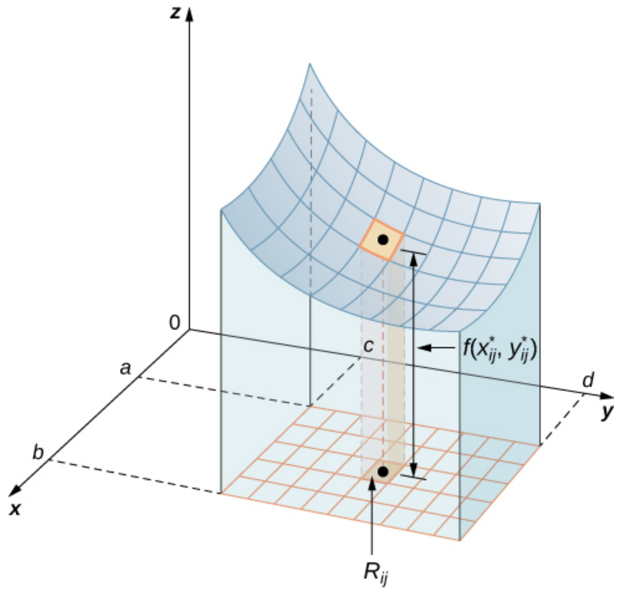
Definizione 2.8.1 (Integrale Doppio). *Una funzione $f : R \to \mathbb{R}$ limitata su R dove $R = [a, b] \times [c, d]$. La funzione f si dice integrabile secondo Riemann su R se*

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

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

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

Se si ha che la funzione $f(x, y) \geq 0$ per ogni $(x, y) \in R$, l'integrale doppio rappresenta il volume della regione solida compresa tra la superficie $z = f(x, y)$ e il piano xy sopra il rettangolo R, ossia un cilindroide curvilineo.



Ogni addendo nelle somme inferiore e superiori è il volume di un parallelepiped di base il rettangolo r_{ij} e altezza rispettivamente $m_{i,j}$ e $M_{i,j}$.

Teorema 2.8.1 (Teorema sulla continuità). *Sia R un rettangolo chiuso e limitato in $\mathbb{R}^2$ definito come $R = [a, b] \times [c, d]$ e sia $f : R \to \mathbb{R}$ una funzione a più variabili. Se la funzione è continua su tutto il rettangolo R, allora f è integrabile secondo Riemann su R.*

2.8.2 Calcolo Integrale sui rettangoli

Sia $f : R \to \mathbb{R}$ una funzione continua definita su un rettangolo $R = [a, b] \times [c, d]$, se $f(x, y) = g(x) \cdot h(y)$ con $g : [a, b] \to \mathbb{R}$ e $h : [c, d] \to \mathbb{R}$ integrabili su i rispettivi intervalli, allora f è integrabile su R e vale la seguente proprietà di separazione:

$$\iint_R f(x, y)\, dx\, dy = \left(\int_a^b g(x)\, dx \right) \cdot \left(\int_c^d h(y)\, dy \right)$$

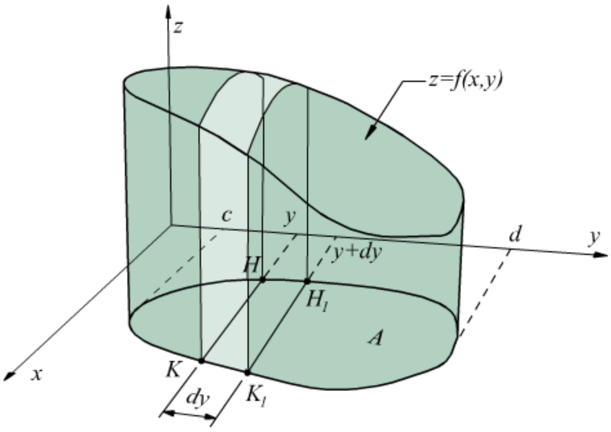
Questo è il caso semplice in cui la funzione si può "separare" in due funzioni di una sola variabile.

Teorema di Fubini-Tonelli

Nel caso generale, si definisce il teorema di Fubini-Tonelli, che permette di calcolare l'integrale doppio come integrali singoli successivi (integrali iterati). Sia $f : R \to \mathbb{R}$ 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).



2.8.3 Integrali doppi su domini generali e la misurabilità secondo Peano-Jordan

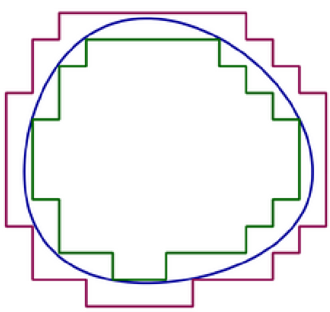
Sia D una regione del piano limitata e sia f definita e limitata su questo dominio. Estendiamo la funzione f a tutto il rettangolo R che contiene D, ponendo $f(x, y) = 0$ per ogni punto $(x, y) \notin D$.

Definizione 2.8.2. *Sia f una funzione definita e limitata su un dominio limitato $D \subset \mathbb{R}^2$. La funzione f si dice integrabile secondo Riemann su D se la funzione estesa $\tilde{f}$ è integrabile secondo Riemann sul rettangolo R che contiene D. In tal caso, l'integrale doppio di f su D è definito come:*

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

Non tutti gli insiemi possono essere misurabili perché non siamo in grado di dire con certezza quanto è grande la sua area. E se non conosciamo l'area di base non possiamo calcolare il volume che si sta sopra. Un sottoinsieme limitato $D \subset \mathbb{R}^2$ è <u>misurabile secondo Peano-Jordan</u> se la funzione f (identicamente 1 su D e 0 fuori) è integrabile secondo Riemann sul rettangolo che contiene D. In tale caso

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



L'idea di misurabilità secondo Peano-Jordan: se riesco a "riempire" l'insieme con dei rettangoli (o a "svuotarlo" con dei rettangoli) in modo che le somme inferiori e superiori coincidano, allora posso definire la sua area.
In generale, ogni rettangolo $R = [a, b] \times [c, d]$ è misurabile secondo Peano-Jordan, e la sua misura è data da:

$$|R| = (b - a)(d - c) \quad \text{che è l'area del rettangolo.}$$

Tutti gli insiemi che conosciamo dalla geometria elementare (poligoni, cerchi, ecc.) sono misurabili secondo Peano-Jordan e la loro misura corrisponde alla formula delle loro aree.

Un esempio di un poligono che non è misurabile secondo Peano-Jordan

Sia Q un quadrato $[1, 0] \times [0, 1]$ con coordinate razionali e sia una funzione f : $Q \to \mathbb{R}$ 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.

2.8.4 Proprietà degl integrali doppi

Analogamente agli integrali singoli:

- **Linearità:** L'integrale di una combinazione lineare è la combinazione lineare degli integrali.

$$\iint_D (\alpha f + \beta g) = \alpha \iint_D f + \beta \iint_D g$$

- **Additività sul dominio:** Se D viene diviso in due sottodomini D_1 e D_2 che non si sovrappongono (se non sul bordo), l'integrale su D è la somma degli integrali.
- **Monotonia:** Se $f(x, y) \leq g(x, y)$ su tutto D, allora $\iint_D f \leq \iint_D g$.
- **Relazione della funzione con il suo valore assoluto:** Se f è integrabile su D, allora $|f|$ è integrabile su D e vale:

$$\left| \iint_D f(x, y)\, dx\, dy \right| \leq \iint_D |f(x, y)|\, dx\, dy$$

E se D è misurabile secondo Peano-Jordan, allora $|D| = \iint_D 1\, dx\, dy$ e posso così definire:

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

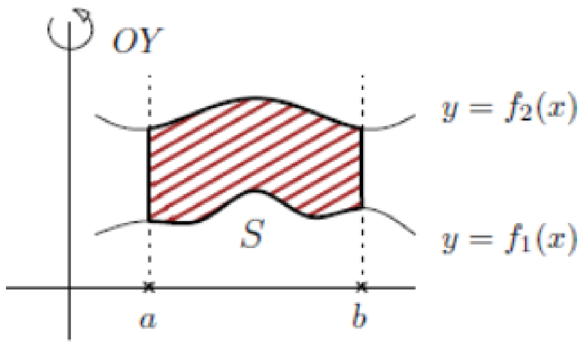
2.8.5 Calcolo Integrale su domini semplici

Non tutti i domini di integrazione sono rettangoli. Spesso ci troviamo a dover integrare su figure dove gli estremi dipendono dalla posizione. Si dicono domini normali (o semplici) quelle regioni del piano delimitate da due rette parallele e dai grafici di due funzioni continue. Ne esistono di due tipi:

- Dominio normale rispetto all'asse x **(Y-semplice):**

  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 $f_2(x)$.

$$D = \{(x, y) \in \mathbb{R}^2 : a \leq x \leq b, \quad f_1(x) \leq y \leq f_2(x)\}$$



  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:

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

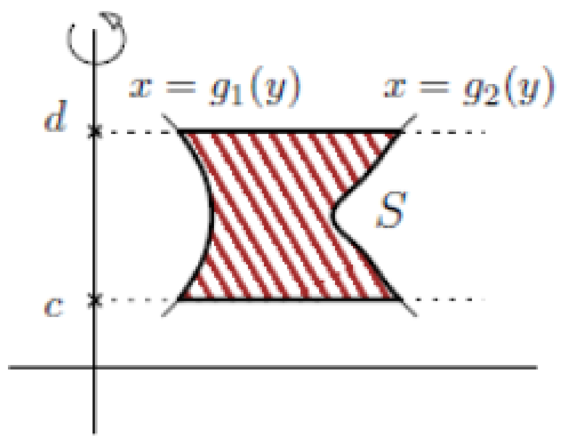
- Dominio normale rispetto all'asse y **(X-semplice):**

  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 $h_2(y)$.

$$D = \{(x, y) \in \mathbb{R}^2 : c \leq y \leq d, \quad h_1(y) \leq x \leq h_2(y)\}$$

  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 insimemi, pensando all'interpretazione geometrica, sono l'area delle regioni che si ottengono:

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

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

Teorema 2.8.2 (Formule di riduzione per domini normali). *Ogni funzione continua $f : D \to \mathbb{R}$ definita su un dominio normale D è integrabile secondo Riemann su D e vale:*

- *Se D è Y-semplice:*

$$\iint_D f(x,y)\, dx\, dy = \int_a^b \left(\int_{f_1(x)}^{f_2(x)} f(x,y)\, dy \right) dx$$

- *Se D è X-semplice:*

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

Supponiamo di avere $D_1, D_2, \ldots, D_k$ 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.

2.8.6 Domini simmetrici e la jacobiana

Nell'analisi unidimensionale il cambio di variabile è possibile se la funzione è di classe C^1 e biunivoca (invertibile):

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

Nell'analisi multidimensionale, spesso quando abbiamo domini che presentano delle simmetrie ci conviene cambiare le variabili trasformando il dominio D (nel piano xy) in un dominio D' (nel piano ρ, θ) molto più semplice (ad esempio un rettangolo), tramite una trasformazione biunivoca.
Sia $f : \mathbb{R}^2 \to \mathbb{R}^2$ tale che

$$(\rho, \theta) \to f(\rho, \theta) = (f_1(\rho, \theta), f_2(\rho, \theta)) = (x(\rho, \theta), y(\rho, \theta))$$

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 <u>Jacobiana</u> della trasformazione:

$$J_f(\rho, \theta) = \begin{bmatrix} \nabla f_1(\rho, \theta) \\ \nabla f_2(\rho, \theta) \end{bmatrix} = \begin{bmatrix} \frac{\partial x}{\partial \rho} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial \rho} & \frac{\partial y}{\partial \theta} \end{bmatrix} = \begin{bmatrix} \cos\theta & -\rho \sin\theta \\ \sin\theta & \rho \cos\theta \end{bmatrix}$$

Il determinante Jacobiano è quindi:

$$\det(J_f(\rho, \theta)) = \rho \cos^2\theta - (-\rho \sin^2\theta) = \rho(\cos^2\theta + \sin^2\theta) = \rho$$

(per la proprietà trigonometrica fondamentale $\cos^2\theta + \sin^2\theta = 1$).
Se il determinante è diverso da zero, la trasformazione è localmente invertibile.
Con $\rho = 0 \implies x = 0,\ y = 0$ (l'origine), la trasformazione non è biunivoca. Inoltre, data la periodicità di 2π, per renderla biunivoca restringiamo il dominio di θ a $[0, 2\pi)$ (o $(-\pi, \pi]$) e $\rho > 0$.
Consideriamo quindi l'aperto $A = (0, +\infty) \times (0, 2\pi)$.

Teorema 2.8.3 (Cambiamento di variabili). *Siano D e T due insiemi aperti misurabili del piano. Sia $F : T \to D$ una trasformazione di coordinate (cambiamento di variabili) data da $F(\rho, \theta) = (x(\rho, \theta), y(\rho, \theta))$.*
Supponiamo che:

- *F sia biunivoca e di classe C^1 sull'interno di T;*
- *Il determinante Jacobiano $\det J_{F(\rho,\theta)}$ sia diverso da zero ovunque in T.*

Allora, per ogni funzione $f(x, y)$ integrabile su D, vale la formula:

$$\iint_D f(x, y)\, dx\, dy = \iint_T f(\rho \cos\theta, \rho \sin\theta) \cdot \rho\, d\rho\, d\theta$$