

Analisi Matematica 1B - a.a. 2020–2021 - Lezione 13

Teoremi sulla differenziabilità

Damiano Foschi

(versione aggiornata il 12 aprile 2021)

Nella precedente lezione abbiamo definito le derivate parziali di una funzione scalare di più variabili,

$$\frac{\partial f}{\partial x_k}(\mathbf{x}) := \lim_{t \rightarrow 0} \frac{f(\mathbf{x} + t\mathbf{e}_k) - f(\mathbf{x})}{t};$$

e più in generale le derivate direzionali rispetto a qualsiasi vettore,

$$D_{\mathbf{v}}f(\mathbf{x}) := \lim_{t \rightarrow 0} \frac{f(\mathbf{x} + t\mathbf{v}) - f(\mathbf{x})}{t}.$$

Si tratta di limiti di rapporti incrementali fatti rispetto ad una variabile reale. Il calcolo di queste derivate si riconduce facilmente al calcolo di derivate di funzioni di una variabile. Diversamente da quanto succede nel caso di funzioni di una variabile, l'esistenza delle derivate parziali o delle derivate direzionali non è sufficiente a garantire la possibilità di avere approssimazioni locali del primo ordine (lineari). Serve una condizione più forte, la differenziabilità. Una funzione è differenziabile in un punto $\mathbf{x} \in \mathbb{R}^n$ quando

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

dove $\nabla f = (\partial_{x_1}f, \dots, \partial_{x_n}f)$ indica il gradiente di f . La verifica della differenziabilità in base a questa definizione equivale alla verifica della validità del limite

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

Si tratta di un limite in n variabili, e la sua verifica può risultare piuttosto laboriosa.

1 Teorema del differenziale totale

La semplice esistenza delle derivate parziali è una condizione necessaria, ma non sufficiente per garantire la differenziabilità della funzione; non è sufficiente nemmeno a garantire la continuità della funzione. Otteniamo però una condizione sufficiente quando le derivate parziali esistono e sono continue, allora risulta che la funzione è sicuramente differenziabile. Questo è il contenuto del teorema del *differenziale totale* che andiamo ad enunciare e a dimostrare.

Teorema 1.1. *Sia $f(\mathbf{x})$ una funzione scalare definita in un intorno del punto \mathbf{p} . Supponiamo che anche le sue derivate parziali siano ben definite nell'intorno di \mathbf{p} . Se tutte le derivate parziali sono continue nel punto \mathbf{p} allora f è differenziabile nel punto \mathbf{p} .*

Dimostrazione. Diamo la dimostrazione nel caso di una funzione di due variabili $f(x, y)$ definita nell'intorno del punto $\mathbf{p} = (x_*, y_*)$. Come ipotesi abbiamo che le derivate parziali $\partial_x f(x, y)$ e $\partial_y f(x, y)$ sono definite in un intorno di (x_*, y_*) e continue in tale punto; questo significa che per ogni $\varepsilon > 0$ esiste $\delta_\varepsilon > 0$ tale che

$$\|(x, y) - (x_*, y_*)\| < \delta_\varepsilon \implies \begin{cases} |\partial_x f(x, y) - \partial_x f(x_*, y_*)| < \varepsilon, \\ |\partial_y f(x, y) - \partial_y f(x_*, y_*)| < \varepsilon. \end{cases} \quad (1)$$

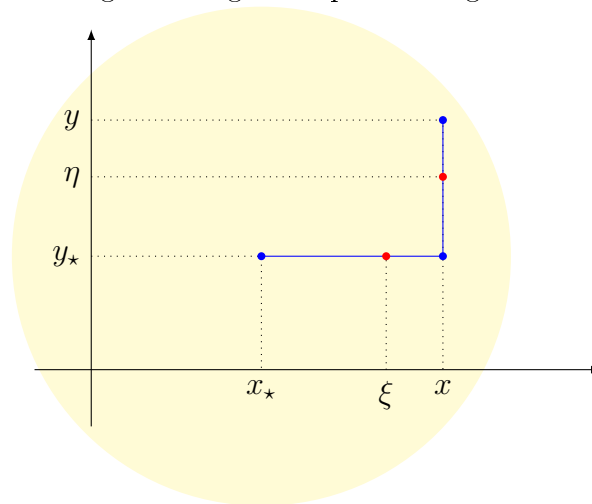
La tesi da dimostrare è che f è differenziabile in (x_*, y_*) , ovvero che vale il limite

$$\lim_{(x, y) \rightarrow (x_*, y_*)} \frac{f(x, y) - f(x_*, y_*) - \partial_x f(x_*, y_*)(x - x_*) - \partial_y f(x_*, y_*)(y - y_*)}{\|(x, y) - (x_*, y_*)\|} = 0. \quad (2)$$

Fissato $\varepsilon > 0$, scegliamo un punto (x, y) nell'intorno sferico di (x_*, y_*) con raggio δ_ε ,

$$\|(x, y) - (x_*, y_*)\| < \delta_\varepsilon. \quad (3)$$

Consideriamo un percorso che parte dal punto (x_*, y_*) , passa dal punto (x, y_*) , e arriva al punto (x, y) , muovendosi lungo due segmenti paralleli agli assi.



Lungo il primo segmento consideriamo la funzione $\varphi(t) := f(t, y_*)$; essa è derivabile per t che varia tra x_* e x , con derivata $\varphi'(t) = \partial_x f(t, y_*)$. Per il teorema di Lagrange, esiste ξ compreso tra x_* e x tale che

$$\frac{f(x, y_*) - f(x_*, y_*)}{x - x_*} = \frac{\varphi(x) - \varphi(x_*)}{x - x_*} = \varphi'(\xi) = \partial_x f(\xi, y_*),$$

e dunque

$$f(x, y_*) - f(x_*, y_*) = \partial_x f(\xi, y_*)(x - x_*). \quad (4)$$

Lungo il secondo segmento consideriamo la funzione $\psi(s) := f(x, s)$; essa è derivabile per s che varia tra y_* e y , con derivata $\psi'(s) = \partial_y f(x, s)$. Per il teorema di Lagrange, esiste η compreso tra y_* e y tale che

$$\frac{f(x, y) - f(x, y_*)}{y - y_*} = \frac{\psi(y) - \psi(y_*)}{y - y_*} = \psi'(\eta) = \partial_y f(x, \eta),$$

e dunque

$$f(x, y) - f(x, y_*) = \partial_y f(x, \eta)(y - y_*). \quad (5)$$

Sommando i membri a destra e a sinistra delle due uguaglianze (5) e (4) otteniamo

$$\begin{aligned} f(x, y) - f(x_*, y_*) &= f(x, y) - f(x, y_*) + f(x, y_*) - f(x_*, y_*) = \\ &= \partial_y f(x, \eta)(y - y_*) + \partial_x f(\xi, y_*)(x - x_*). \end{aligned}$$

Sottraiamo ora, a destra e a sinistra, la quantità $\partial_x f(x_*, y_*)(x - x_*) + \partial_y f(x_*, y_*)(y - y_*)$, e otteniamo la quantità che si trova al numeratore nel limite (2),

$$\begin{aligned} f(x, y) - f(x_*, y_*) - \partial_x f(x_*, y_*)(x - x_*) - \partial_y f(x_*, y_*)(y - y_*) &= \\ = (\partial_x f(\xi, y_*) - \partial_x f(x_*, y_*))(x - x_*) + (\partial_y f(x, \eta) - \partial_y f(x_*, y_*))(y - y_*). \end{aligned}$$

I punti (ξ, y_*) e (x, η) sono più vicini a (x_*, y_*) di quanto lo sia (x, y) , e dunque abbiamo

$$\|(\xi, y_*) - (x_*, y_*)\| < \delta_\varepsilon, \quad \|(x, \eta) - (x_*, y_*)\| < \delta_\varepsilon,$$

per l'ipotesi di continuità (1) delle derivate parziali segue che

$$|\partial_x f(\xi, y_*) - \partial_x f(x_*, y_*)| < \varepsilon, \quad |\partial_y f(x, \eta) - \partial_y f(x_*, y_*)| < \varepsilon.$$

Inoltre $|x - x_*|$ e $|y - y_*|$ sono entrambi minori di $\|(x, y) - (x_*, y_*)\|$. Otteniamo allora

$$\begin{aligned} |f(x, y) - f(x_*, y_*) - \partial_x f(x_*, y_*)(x - x_*) - \partial_y f(x_*, y_*)(y - y_*)| &\leq \\ \leq |\partial_x f(\xi, y_*) - \partial_x f(x_*, y_*)| |x - x_*| + |\partial_y f(x, \eta) - \partial_y f(x_*, y_*)| |y - y_*| &\leq \\ \leq 2\varepsilon \|(x, y) - (x_*, y_*)\|, \end{aligned}$$

ovvero

$$\left| \frac{f(x, y) - f(x_*, y_*) - \partial_x f(x_*, y_*)(x - x_*) - \partial_y f(x_*, y_*)(y - y_*)}{\|(x, y) - (x_*, y_*)\|} \right| \leq 2\varepsilon. \quad (6)$$

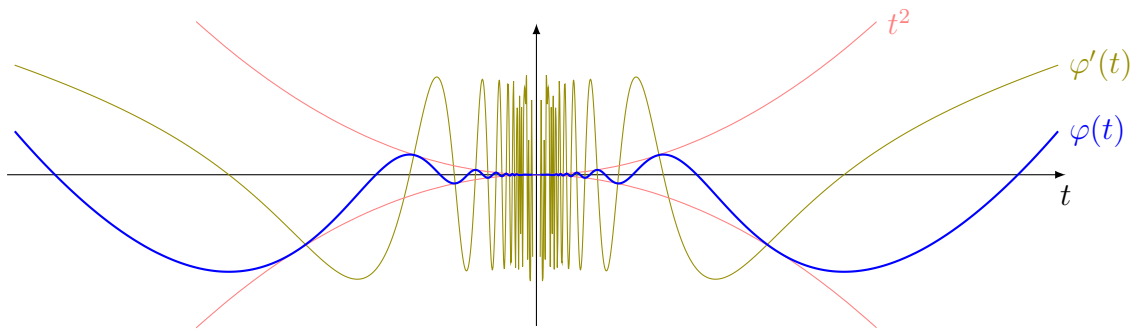
E tutto questo a partire dalla scelta (3) di (x, y) nell'intorno di (x_*, y_*) di raggio δ_ε . Essendo ε una qualsiasi quantità positiva, ciò significa che la quantità in (6) è infinitesima per (x, y) che tende a (x_*, y_*) , e dunque vale il limite (2). \square

Esempio 1.2. Alla luce del teorema 1.1 riprendiamo un esempio che abbiamo già visto nella precedente lezione. Consideriamo la funzione $f(x, y) := x \sin(y)$ e verifichiamo se f è differenziabile nel punto $(-2, \pi/3)$. Le derivate parziali di f sono $\partial_x f(x, y) = \sin(y)$ e $\partial_y f(x, y) = x \cos(y)$, e siccome sono funzioni continue in ogni punto per il teorema del differenziale totale la funzione f è differenziabile in ogni punto. Ora ci è bastata una riga per verificarlo, mentre nella scorsa lezione ci sono volute un paio di pagine.

Il teorema del differenziale totale ci fornisce una condizione sufficiente per la differenziabilità. La continuità delle derivate parziali non è comunque una condizione necessaria, ci possono essere funzioni differenziabili con derivate parziali non continue.

Esempio 1.3. Consideriamo la funzione di una variabile definita da

$$\varphi(t) := \begin{cases} t^2 \cos \frac{1}{t}, & \text{se } t \neq 0, \\ 0, & \text{se } t = 0. \end{cases}$$



Quando $t \neq 0$ abbiamo $\varphi'(t) = 2t \cos \frac{1}{t} + \sin \frac{1}{t}$; mentre per $t = 0$ abbiamo

$$\varphi'(0) = \lim_{h \rightarrow 0} \frac{\varphi(h) - \varphi(0)}{h} = \lim_{h \rightarrow 0} h \cos \frac{1}{h} = 0.$$

Dunque la derivata $\varphi'(t)$ è definita per ogni $t \in \mathbb{R}$; ma non è continua in $t = 0$, infatti il limite $\lim_{t \rightarrow 0} \varphi'(t)$ non esiste, in quanto non esiste il limite di $\sin \frac{1}{t}$.

Con questa funzione φ costruiamo una funzione di due variabili ponendo

$$f(x, y) = \varphi(x) + \varphi(y), \quad \forall (x, y) \in \mathbb{R}^2.$$

Le derivate parziali sono definite in ogni punto,

$$\partial_x f(x, y) = \varphi'(x), \quad \partial_y f(x, y) = \varphi'(y),$$

ma non sono continue nell'origine. Nell'origine si annullano entrambe e dunque il gradiente è nullo, $\nabla f(0, 0) = (0, 0)$. Anche la funzione si annulla nell'origine, $f(0, 0) = 0$. La condizione di differenziabilità nell'origine allora si riduce a richiedere che

$$f(x, y) = o(\|(x, y)\|), \quad \text{per } (x, y) \rightarrow (0, 0).$$

Controlliamo se tale condizione è verificata. Siccome $|\varphi(t)| \leq t^2$ abbiamo che

$$|f(x, y)| \leq |\varphi(x)| + |\varphi(y)| \leq x^2 + y^2 = \|(x, y)\|^2,$$

e dunque

$$\left| \frac{f(x, y)}{\|(x, y)\|} \right| \leq \|(x, y)\| \rightarrow 0, \quad \text{per } (x, y) \rightarrow (0, 0).$$

Dunque f è effettivamente differenziabile in $(0, 0)$, pur non avendo derivate parziali continue.

Esercizio 1.4. Studia la derivabilità e la differenziabilità delle seguenti funzioni definite su \mathbb{R}^2 e verifica se le derivate parziali sono continue.

$$A(x, y) := \begin{cases} \frac{x^2 y - x y^2}{x^2 + y^2}, & \text{se } (x, y) \neq (0, 0), \\ 0, & \text{se } (x, y) = (0, 0); \end{cases} \quad B(x, y) := \begin{cases} \frac{\sin \sqrt{x^2 + y^2}}{\sqrt{x^2 + y^2}}, & \text{se } (x, y) \neq (0, 0), \\ 1, & \text{se } (x, y) = (0, 0); \end{cases}$$

$$C(x, y) := \begin{cases} \frac{1 - \cos(xy)}{x^2 + y^2}, & \text{se } (x, y) \neq (0, 0), \\ 0, & \text{se } (x, y) = (0, 0); \end{cases} \quad D(x, y) := \begin{cases} x \log \frac{x^2 + 3y^2}{x^2 + y^2}, & \text{se } (x, y) \neq (0, 0), \\ 0, & \text{se } (x, y) = (0, 0). \end{cases}$$

Definizione 1.5. Una funzione $f: E \rightarrow \mathbb{R}$ definita sul dominio $E \subset \mathbb{R}^n$ si dice di classe C^1 su E quando esistono le sue derivate parziali su E ed esse sono continue.

Per il teorema del differenziale totale ogni funzione di classe C^1 è anche differenziabile in ogni punto interno del dominio. In particolare ogni funzione di classe C^1 definita su un aperto è sempre differenziabile.

2 Derivate seconde

Ogni derivata parziale di una funzione scalare di più variabili è a sua volta una funzione scalare; possiamo dunque provare a calcolare le sue derivate parziali. La derivata parziale di una derivata parziale di una funzione scalare f si dice *derivata seconda* di f . Per le derivate seconde possiamo usare le notazioni

$$\frac{\partial^2 f}{\partial x_j \partial x_k}(\mathbf{x}), \quad \partial_{x_j} \partial_{x_k} f(\mathbf{x}), \quad f_{x_j x_k}(\mathbf{x}),$$

per indicare la derivata rispetto a x_j della derivata rispetto a x_k di f calcolata nel punto \mathbf{x} .

Esempio 2.1. Consideriamo la funzione $f(x, y) := x^4 - x^3 y + y^2$. Le sue derivate parziali prime sono

$$\frac{\partial f}{\partial x}(x, y) = 4x^3 - 3x^2 y, \quad \frac{\partial f}{\partial y}(x, y) = -x^3 + 2y.$$

Derivando ulteriormente troviamo le derivate parziali seconde

$$\begin{aligned}\frac{\partial^2 f}{\partial x \partial x}(x, y) &= 12x^2 - 6xy, & \frac{\partial^2 f}{\partial y \partial x}(x, y) &= -3x^2, \\ \frac{\partial^2 f}{\partial x \partial y}(x, y) &= -3x^2, & \frac{\partial^2 f}{\partial y \partial y}(x, y) &= 2.\end{aligned}$$

Osserviamo che in questo esempio le derivate seconde miste $\partial_y \partial_x f$ e $\partial_x \partial_y f$ coincidono.

Esempio 2.2. Calcoliamo le derivate seconde miste $\partial_y \partial_x f(0, 0)$ e $\partial_x \partial_y f(0, 0)$ per la funzione definita da

$$f(x, y) := \begin{cases} \frac{x^3 y - x y^3}{x^2 + y^2}, & \text{se } (x, y) \neq (0, 0), \\ 0, & \text{se } (x, y) = (0, 0). \end{cases}$$

Prima ci serve di conoscere le derivate prime non solo nell'origine, ma anche in un intorno dell'origine. Siccome $f(x, 0) = 0$ per ogni x abbiamo che $\partial_x f(0, 0) = 0$. Essendo f una funzione razionale, nei punti in cui non si annulla il denominatore possiamo calcolare le derivate con le solite regole di derivazione, dunque per $(x, y) \neq (0, 0)$ abbiamo

$$\partial_x f(x, y) = \frac{(3x^2 y - y^3)(x^2 + y^2) - (x^3 y - x y^3)2x}{(x^2 + y^2)^2} = \frac{x^4 y + 4x^2 y^3 - y^5}{(x^2 + y^2)^2}. \quad (7)$$

Siccome $\partial_x f(0, y) = \frac{0 + 0 - y^5}{(0 + y^2)^2} = -y$, derivando rispetto ad y otteniamo $\partial_y \partial_x f(0, 0) = -1$. Scambiando x con y la funzione cambia di segno, $f(x, y) = -f(y, x)$, e per questa simmetria risulta che $\partial_y f(x, y) = -\partial_x f(y, x)$ e anche $\partial_x \partial_y f(x, y) = -\partial_y \partial_x f(y, x)$. In particolare abbiamo $\partial_x \partial_y f(0, 0) = -\partial_y \partial_x f(0, 0) = +1$. Dunque derivate seconde miste rispetto alla stessa coppia di variabili nello stesso punto possono assumere valore differente a seconda dell'ordine con cui si deriva. Osserviamo che per $(x, y) \neq (0, 0)$ derivando l'espressione (7) si ottiene

$$\partial_y \partial_x f(x, y) = \frac{x^6 + 9x^4 y^2 - 9x^2 y^4 - y^6}{(x^2 + y^2)^3},$$

in particolare per $x \neq 0$ abbiamo $\partial_y \partial_x f(x, 0) = \frac{x^6}{x^6} = +1$, e questo ci fa capire che $\partial_y \partial_x f$ non può essere continua in $(0, 0)$.

Il teorema di Schwarz ci assicura che quando le derivate seconde miste sono continue l'ordine con cui si effettuano le derivate non influisce sul risultato finale.

Teorema 2.3 (Schwarz). *Sia $f(\mathbf{x})$ una funzione definita in un intorno del punto $\mathbf{p} \in \mathbb{R}^n$. Supponiamo che siano definite in un intorno di \mathbf{p} anche le derivate parziali*

$$\partial_{x_j} f, \quad \partial_{x_k} f, \quad \partial_{x_j} \partial_{x_k} f, \quad \partial_{x_k} \partial_{x_j} f,$$

per due diversi indici j e k compresi tra 1 e n . Se le derivate seconde miste $\partial_{x_j} \partial_{x_k} f$ e $\partial_{x_k} \partial_{x_j} f$ sono continue nel punto \mathbf{p} allora i loro valori coincidono,

$$\partial_{x_j} \partial_{x_k} f(\mathbf{p}) = \partial_{x_k} \partial_{x_j} f(\mathbf{p}).$$

Come per la dimostrazione del teorema del differenziale totale (teorema (1.1)) anche nella dimostrazione del teorema di Schwarz faremo un uso ripetuto del teorema di Lagrange.

Dimostrazione. È sufficiente considerare il caso di una funzione di due variabili $f(x, y)$ definita e derivabile due volte intorno al punto $\mathbf{p} = (x_*, y_*)$. Fissato (x, y) in un intorno di (x_*, y_*) definiamo le funzioni

$$A(t) := f(t, y) - f(t, y_*), \quad B(s) := f(x, s) - f(x_*, s).$$

Si tratta di funzioni derivabili. Applicando il teorema di Lagrange alla funzione $A(t)$ nell'intervallo di estremi x_* e x troviamo che esiste un punto ξ_1 in tale intervallo tale che

$$\begin{aligned} \frac{(f(x, y) - f(x, y_*)) - (f(x_*, y) - f(x_*, y_*))}{x - x_*} &= \\ &= \frac{A(x) - A(x_*)}{x - x_*} = A'(\xi_1) = \partial_x f(\xi_1, y) - \partial_x f(\xi_1, y_*). \end{aligned} \quad (8)$$

Applicando di nuovo il teorema di Lagrange alla funzione $s \mapsto \partial_x f(\xi_1, s)$ nell'intervallo di estremi y_* e y troviamo che esiste un η_1 in tale intervallo tale che

$$\frac{\partial_x f(\xi_1, y) - \partial_x f(\xi_1, y_*)}{y - y_*} = \partial_y \partial_x f(\xi_1, \eta_1). \quad (9)$$

Combinando le formule (8) e (9) ricaviamo che

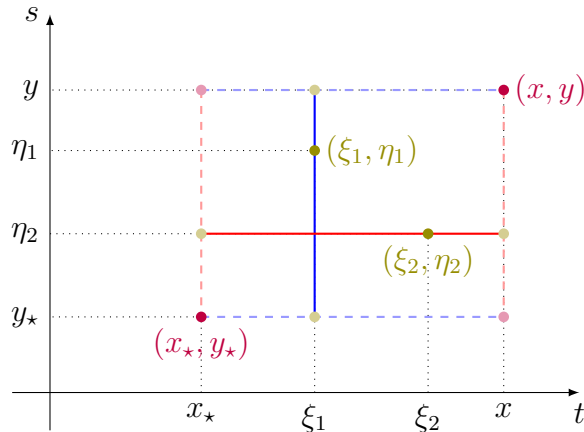
$$f(x, y) - f(x, y_*) - f(x_*, y) + f(x_*, y_*) = \partial_y \partial_x f(\xi_1, \eta_1)(x - x_*)(y - y_*). \quad (10)$$

Procedendo esattamente nello stesso modo per la funzione $B(s)$ invertendo i ruoli delle variabili x e y , si ottiene che esistono un punto η_2 , compreso tra y_* e y , e un punto ξ_2 , compreso tra x_* e x , tali che

$$f(x, y) - f(x_*, y) - f(x, y_*) + f(x_*, y_*) = \partial_x \partial_y f(\xi_2, \eta_2)(x - x_*)(y - y_*). \quad (11)$$

Confrontando (10) con (11) ricaviamo che

$$\partial_y \partial_x f(\xi_1, \eta_1) = \partial_x \partial_y f(\xi_2, \eta_2) \quad (12)$$



La scelta dei punti (ξ_1, η_1) e (ξ_2, η_2) dipende dalla scelta del punto iniziale (x, y) ; inoltre, la distanza di (ξ_1, η_1) e (ξ_2, η_2) dal punto (x_*, y_*) è minore della distanza di (x, y) da (x_*, y_*) . Questo significa che quando facciamo tendere (x, y) a (x_*, y_*) anche (ξ_1, η_1) e (ξ_2, η_2) convergono a (x_*, y_*) . Per l'ipotesi di continuità delle derivate $\partial_y \partial_x f$ e $\partial_x \partial_y f$ in (x_*, y_*) , passando al limite per $(x, y) \rightarrow (x_*, y_*)$ dall'uguaglianza (12) otteniamo il risultato cercato,

$$\partial_y \partial_x f(x_*, y_*) = \partial_x \partial_y f(x_*, y_*).$$

□

3 Matrice Hessiana

Una funzione scalare di n variabili può avere n derivate parziali, e ciascuna delle sue derivate prime può avere a sua volta n derivate parziali; dunque ogni funzione scalare può avere $n \cdot n = n^2$ derivate seconde. Gli n^2 valori delle derivate seconde possono essere organizzate in modo da formare una matrice $n \times n$.

Definizione 3.1. Data una funzione scalare $f(\mathbf{x})$ di n variabili che ammette tutte le derivate seconde nel punto \mathbf{p} definiamo *matrice Hessiana* per la funzione f nel punto \mathbf{p} la matrice quadrata $n \times n$ formata dai valori delle derivate parziali seconde di f ,

$$H_f(\mathbf{p}) := \begin{pmatrix} \partial_{x_1} \partial_{x_1} f(\mathbf{p}) & \partial_{x_1} \partial_{x_2} f(\mathbf{p}) & \dots & \partial_{x_1} \partial_{x_n} f(\mathbf{p}) \\ \partial_{x_2} \partial_{x_1} f(\mathbf{p}) & \partial_{x_2} \partial_{x_2} f(\mathbf{p}) & \dots & \partial_{x_2} \partial_{x_n} f(\mathbf{p}) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_{x_n} \partial_{x_1} f(\mathbf{p}) & \partial_{x_n} \partial_{x_2} f(\mathbf{p}) & \dots & \partial_{x_n} \partial_{x_n} f(\mathbf{p}) \end{pmatrix}.$$

Dunque, $(H_f(\mathbf{p}))_{ij} = \partial_{x_i} \partial_{x_j} f(\mathbf{p})$ per $i, j = 1, \dots, n$.

Esempio 3.2. Sia $f(x, y, z) = xe^{2y} \sin(3z)$. Il gradiente di f è

$$\nabla f(x, y, z) = (e^{2y} \sin(3z), 2xe^{2y} \sin(3z), 3xe^{2y} \cos(3z)).$$

La matrice Hessiana è

$$H_f(x, y, z) = \begin{pmatrix} 0, & 2e^{2y} \sin(3z), & 3e^{2y} \cos(3z) \\ 2e^{2y} \sin(3z), & 4xe^{2y} \sin(3z), & 6xe^{2y} \cos(3z) \\ 3e^{2y} \cos(3z), & 6xe^{2y} \cos(3z), & -9xe^{2y} \sin(3z) \end{pmatrix}.$$

Nel punto $(1, 0, \pi/2)$ abbiamo $\nabla f(1, 0, \pi/2) = (-1, -2, 0)$ e

$$H_f\left(1, 0, \frac{\pi}{2}\right) = \begin{pmatrix} 0 & -2 & 0 \\ -2 & -4 & 0 \\ 0 & 0 & 9 \end{pmatrix}.$$

Esempio 3.3. Consideriamo un generico polinomio di secondo grado nelle due variabili x e y ,

$$P(x, y) = a_0 + a_1x + a_2y + a_3x^2 + a_4xy + a_5y^2.$$

Calcoliamo le sue derivate

$$\begin{aligned} P_x(x, y) &= a_1 + 2a_3x + a_4y, & P_y(x, y) &= a_2 + a_4x + 2a_5y, \\ P_{xx}(x, y) &= 2a_3, & P_{xy}(x, y) &= P_{yx}(x, y) = a_4, & P_{yy}(x, y) &= 2a_5. \end{aligned}$$

Dunque,

$$\nabla P(0, 0) = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \quad H_P(0, 0) = \begin{pmatrix} 2a_3 & a_4 \\ a_4 & 2a_5 \end{pmatrix}$$

e risulta che possiamo scrivere

$$P(x, y) = P(0, 0) + \nabla P(0, 0) \cdot \begin{pmatrix} x \\ y \end{pmatrix} + \frac{1}{2} \begin{pmatrix} x \\ y \end{pmatrix} \cdot H_P(0, 0) \begin{pmatrix} x \\ y \end{pmatrix}.$$

Esempio 3.4. Calcoliamo le derivate seconde della norma. Sia $f(\mathbf{x}) := \|\mathbf{x}\|$ per $\mathbf{x} \in \mathbb{R}^n$. Abbiamo già calcolato nella precedente lezione le derivate prime per $\mathbf{x} \neq \mathbf{0}$,

$$\partial_{x_j} f(\mathbf{x}) = \frac{x_j}{\|\mathbf{x}\|}, \quad j = 1, \dots, n.$$

Deriviamo ulteriormente rispetto a x_i ,

$$\begin{aligned} \partial_{x_i} \partial_{x_j} f(\mathbf{x}) &= \partial_{x_i} (x_j \|\mathbf{x}\|^{-1}) \\ &= (\partial_{x_i} x_j) \|\mathbf{x}\|^{-1} + x_j (\partial_{x_i} \|\mathbf{x}\|^{-1}) \\ &= \delta_{ij} \|\mathbf{x}\|^{-1} - x_j \|\mathbf{x}\|^{-2} (\partial_{x_i} \|\mathbf{x}\|) \\ &= \delta_{ij} \|\mathbf{x}\|^{-1} - x_i x_j \|\mathbf{x}\|^{-3} \\ &= \|\mathbf{x}\|^{-3} (\delta_{ij} \|\mathbf{x}\|^2 - x_i x_j), \end{aligned}$$

dove $\delta_{ij} = 0$ quando $i \neq j$ e $\delta_{ij} = 1$ quando $i = j$. Nel caso tridimensionale, $n = 3$, abbiamo

$$H_{\|\cdot\|}(x, y, z) = (x^2 + y^2 + z^2)^{-\frac{3}{2}} \begin{pmatrix} y^2 + z^2 & -xy & -xz \\ -xy & x^2 + z^2 & -yz \\ -xz & -yz & x^2 + y^2 \end{pmatrix}.$$

Esercizio 3.5. Calcola le derivate seconde e matrici Hessiane per le seguenti funzioni:

$$\begin{array}{lll} \log(x + y^2), & \arctan \frac{x + y}{1 - xy}, & x^y, \\ \sqrt{x + yz}, & \frac{xyz}{x + y + z}, & \sin(x \sin(y \sin(z))). \end{array}$$

Definizione 3.6. Una funzione $f: E \rightarrow \mathbb{R}$ definita sul dominio $E \subset \mathbb{R}^n$ si dice di *classe C^2* su E quando esistono le sue derivate parziali prime e seconde su E ed esse sono continue.

Per il teorema del differenziale totale le derivate parziali prime di una funzione di classe C^2 sono differenziabili nei punti interni del dominio. Inoltre, come conseguenza del teorema di Schwarz la matrice Hessiana di ogni funzione di classe C^2 è sempre una matrice simmetrica. Questo fatto è una cosa importante perché sappiamo (dall'insegnamento di Algebra e Geometria) che ogni matrice simmetrica è diagonalizzabile, ed esiste una base ortonormale di \mathbb{R}^n formata da autovettori della matrice.

4 Forme quadratiche e polinomi di secondo grado

Definizione 4.1. Una *forma quadratica* su \mathbb{R}^n è un polinomio omogeneo di secondo grado in n variabili.

Per l'omogeneità, tutti i monomi che compongono una forma quadratica sono di secondo grado, quindi essa si può sempre scrivere nella forma

$$Q(\mathbf{x}) = \sum_{i,j=1}^n c_{ij}x_i x_j = \sum_{i=1}^n \left(x_i \sum_{j=1}^n c_{ij}x_j \right), \quad \forall \mathbf{x} = (x_1, \dots, x_n)^t \in \mathbb{R}^n, \quad (13)$$

per opportuni coefficienti numerici c_{ij} che possiamo leggere come le componenti di una matrice quadrata $C = (c_{ij})_{\substack{i=1,\dots,n \\ j=1,\dots,n}}$. Ad ogni matrice quadrata C con n righe e n colonne possiamo quindi associare la forma quadratica Q definita da (13).

Consideriamo \mathbf{x} come vettore colonna, e osserviamo che la i -esima componente del vettore ottenuto dal prodotto matriciale $C\mathbf{x}$ è data dalla somma $\sum_{j=1}^n c_{ij}x_j$; possiamo allora rileggere la formula (13) usando la notazione matriciale o il calcolo vettoriale, e otteniamo una sequenza di formule equivalenti,

$$Q(\mathbf{x}) = \mathbf{x}^t(C\mathbf{x}) = \mathbf{x} \cdot (C\mathbf{x}) = (C\mathbf{x}) \cdot \mathbf{x} = (C\mathbf{x})^t \mathbf{x} = (\mathbf{x}^t C^t) \mathbf{x} = \mathbf{x}^t (C^t \mathbf{x}) = \mathbf{x} \cdot (C^t \mathbf{x}).$$

Da queste formule si deduce che la forma quadratica Q coincide anche con la forma quadratica associata alla matrice trasposta $C^t = (c_{ji})_{\substack{i=1,\dots,n \\ j=1,\dots,n}}$. Sommando insieme le due espressioni $Q(\mathbf{x}) = \mathbf{x} \cdot (C\mathbf{x})$ e $Q(\mathbf{x}) = \mathbf{x} \cdot (C^t \mathbf{x})$ e dividendo per due otteniamo

$$Q(\mathbf{x}) = \frac{1}{2} (\mathbf{x} \cdot C\mathbf{x} + \mathbf{x} \cdot C^t \mathbf{x}) = \mathbf{x} \cdot \left(\frac{1}{2} (C + C^t) \right) \mathbf{x} = \mathbf{x} \cdot \tilde{C} \mathbf{x}. \quad (14)$$

Dunque la forma quadratica Q coincide anche con la forma quadratica associata alla matrice simmetrica $\tilde{C} = \tilde{C}^t$ definita da

$$\tilde{C} = (\tilde{c}_{ji})_{\substack{i=1,\dots,n \\ j=1,\dots,n}} = \frac{1}{2} (C + C^t) = \left(\frac{c_{ij} + c_{ji}}{2} \right)_{\substack{i=1,\dots,n \\ j=1,\dots,n}}.$$

Quando ci riferiamo ad una forma quadratica possiamo quindi sempre supporre che i coefficienti che appaiono nella formula (13) siano simmetrici, $c_{ij} = c_{ji}$. Si verifica facilmente che vi è una corrispondenza biunivoca tra l'insieme delle forme quadratiche su \mathbb{R}^n e l'insieme delle matrici $n \times n$ simmetriche.

Sappiamo (dall'insegnamento di Algebra e Geometria) che ogni matrice simmetrica C è diagonalizzabile, ovvero che esiste una matrice ortogonale U (dunque invertibile con inversa uguale alla sua trasposta, $U^t = U^{-1}$) tale che

$$U^t C U = D,$$

dove D è una matrice diagonale; le colonne della matrice U formano una base ortonormale $B = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ di autovettori di C e gli elementi della diagonale di D sono i corrispondenti autovalori,

$$U = [\mathbf{u}_1, \dots, \mathbf{u}_n], \quad D = \text{diag}(\lambda_1, \dots, \lambda_n), \quad C\mathbf{u}_k = \lambda_k \mathbf{u}_k, \quad \mathbf{u}_j \cdot \mathbf{u}_k = \begin{cases} 1, & \text{se } j = k, \\ 0, & \text{se } j \neq k. \end{cases}$$

Rileggiamo la forma quadratica attraverso coordinate relative alla base B formata dagli autovettori: utilizzando il vettore $\mathbf{y} = (y_1, \dots, y_n)^t$ formato dalle coordinate che descrivono \mathbf{x} rispetto alla base B , ovvero $\mathbf{x} = \sum_{k=1}^n y_k \mathbf{u}_k = U\mathbf{y}$, troviamo che la forma quadratica si può scrivere come una somma algebrica di quadrati,

$$Q(\mathbf{x}) = \mathbf{x}^t C \mathbf{x} = (U\mathbf{y})^t C (U\mathbf{y}) = \mathbf{y}^t (U^t C U) \mathbf{y} = \mathbf{y}^t D \mathbf{y} = \sum_{k=1}^n \lambda_k y_k^2.$$

La matrice C e la forma quadratica Q si dicono:

- *semidefinite positive*, quando gli autovalori sono tutti non negativi, $\lambda_k \geq 0$ per ogni k , e quindi abbiamo $Q(\mathbf{x}) \geq 0$ per ogni $\mathbf{x} \in \mathbb{R}^n$;
- *semidefinite negative*, quando gli autovalori sono tutti non positivi, $\lambda_k \leq 0$ per ogni k , e quindi abbiamo $Q(\mathbf{x}) \leq 0$ per ogni $\mathbf{x} \in \mathbb{R}^n$;
- *definite positive*, quando gli autovalori sono tutti strettamente positivi, $\lambda_k > 0$ per ogni k , e quindi abbiamo $Q(\mathbf{x}) > 0$ per ogni $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$;
- *definite negative*, quando gli autovalori sono tutti strettamente negativi, $\lambda_k < 0$ per ogni k , e quindi abbiamo $Q(\mathbf{x}) < 0$ per ogni $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$;
- *indefinite*, quando esistono autovalori sia positivi che negativi, e quindi $Q(\mathbf{x})$ assume sia valori positivi che valori negativi al variare di $\mathbf{x} \in \mathbb{R}^n$.

Ad esempio:

- $Q(x, y) = 2x^2 + 3y^2$ è definita positiva in quanto $Q(x, y) \geq 0$ per ogni (x, y) e $Q(x, y) = 0$ solo per $(x, y) = (0, 0)$ (gli autovalori sono 2 e 3);
- $Q(x, y) = (x - 2y)^2$ è semidefinita positiva, ma non definita positiva, in quanto $Q(x, y) \geq 0$ per ogni (x, y) , ma $Q(2, 1) = 0$ (gli autovalori sono 0 e 5);
- $Q(x, y) = 4xy$ è indefinita, in quanto $Q(1, 1) = 4 > 0$ e $Q(1, -1) = -4 < 0$ (gli autovalori sono -2 e 2).

Esempio 4.2. Consideriamo la forma quadratica su \mathbb{R}^2 definita da

$$Q(x, y) := x^2 + 4xy + 3y^2.$$

La matrice simmetrica corrispondente a Q è

$$C = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}.$$

Il polinomio caratteristico della matrice è

$$\det(C - \lambda I) = (1 - \lambda)(3 - \lambda) - 2 \cdot 2 = \lambda^2 - 4\lambda - 1.$$

Gli autovalori della matrice sono i valori di λ che annullano il polinomio caratteristico, ovvero $\lambda_1 = 2 - \sqrt{5} < 0$, $\lambda_2 = 2 + \sqrt{5} > 0$. Essendo i due autovalori di segno opposto la forma quadratica risulta essere indefinita.

Esempio 4.3. Consideriamo la forma quadratica su \mathbb{R}^3 definita da

$$Q(x, y, z) := 5x^2 - 4xy - 4xz + 4y^2 + z^2.$$

La matrice simmetrica corrispondente a Q è

$$C = \begin{pmatrix} 5 & -2 & -2 \\ -2 & 4 & 0 \\ -2 & 0 & 1 \end{pmatrix}.$$

Il polinomio caratteristico della matrice è

$$\det(C - \lambda I) = -\lambda^3 + 10\lambda^2 - 21\lambda = -\lambda(\lambda - 3)(\lambda - 7).$$

Gli autovalori della matrice sono dunque $\lambda_1 = 0$, $\lambda_2 = 3$, $\lambda_3 = 7$. Essendo tutti gli autovalori non negativi e uno di essi nullo la forma quadratica risulta essere semidefinita positiva.

Esercizio 4.4. Scrivi le forme quadratiche associate alle seguenti matrici:

$$C_1 := \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}; \quad C_2 := \begin{pmatrix} -1 & 2 \\ -2 & -1 \end{pmatrix}; \quad C_3 := \begin{pmatrix} 0 & 1 & 2 \\ 1 & 2 & 0 \\ 2 & 0 & 1 \end{pmatrix}; \quad C_4 := \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$

Esercizio 4.5. Classifica (ovvero determina se sono definite, o semidefinite, positive o negative, o indefinite) le seguenti forme quadratiche:

$$\begin{aligned} Q_1(x, y) &:= x^2 + (x - y)^2 - y^2; & Q_4(x, y, z) &:= xy + xz + yz; \\ Q_2(x, y) &:= x^2 + 2(x - y)^2 - 2y^2; & Q_5(x, y, z) &:= -x^2 + 2xy - 2y^2 + 2yz - z^2; \\ Q_3(x, y) &:= 3x^2 + 2(x - y)^2 - y^2; & Q_6(x, y, z) &:= 2x^2 + 2xy + 2xz + 2y^2 + 2yz + 2z^2. \end{aligned}$$

Calcoliamo gradiente e matrice Hessiana di una generica forma quadratica

$$Q(\mathbf{x}) = \mathbf{x} \cdot C \mathbf{x} = \sum_{i,j=1}^n c_{ij} x_i x_j,$$

associata alla matrice simmetrica $C = C^t$, $c_{ij} = c_{ji}$. Per aiutarci nel calcolo utilizziamo la notazione della *delta di Kronecker* per indicare le derivate parziali delle coordinate,

$$\delta_{ik} := \partial_{x_k} x_i = \begin{cases} 1, & \text{se } k = i, \\ 0, & \text{se } k \neq i. \end{cases}$$

Le derivate prime di Q sono date da

$$\begin{aligned} \partial_{x_k} Q(\mathbf{x}) &= \sum_{ij} c_{ij} \partial_{x_k} (x_i x_j) = \sum_{ij} c_{ij} (\delta_{ik} x_j + x_i \delta_{jk}) = \\ &= \sum_j c_{kj} x_j + \sum_i c_{ik} x_i = 2 \sum_i c_{ki} x_i = 2(C\mathbf{x})_k, \end{aligned}$$

e dunque il gradiente è dato da

$$\nabla Q(\mathbf{x}) = 2C\mathbf{x}.$$

Le derivate seconde di Q sono date da

$$\partial_{x_j x_k}^2 Q(\mathbf{x}) = \partial_{x_j} \left(2 \sum_i c_{ki} x_i \right) = 2 \sum_i c_{ki} \delta_{ij} = 2c_{kj},$$

e dunque la matrice Hessiana è data da

$$H_Q(\mathbf{x}) = 2C.$$

Un generico polinomio di secondo grado nelle n variabili x_1, x_2, \dots, x_n può sempre essere scritto nella forma

$$P(\mathbf{x}) = a + \sum_{i=1}^n b_i x_i + \sum_{i,j=1}^n c_{ij} x_i x_j = a + \mathbf{b} \cdot \mathbf{x} + \mathbf{x} \cdot (C\mathbf{x}),$$

con coefficienti descritti da:

- un valore scalare $a \in \mathbb{R}$ per il termine noto;
- un vettore $\mathbf{b} = (b_i)_{i=1, \dots, n} \in \mathbb{R}^n$ per i termini di primo grado;
- una matrice quadrata $C = (c_{ij})_{\substack{i=1, \dots, n \\ j=1, \dots, n}}$ di ordine $n \times n$, che per quanto visto prima possiamo assumere simmetrica, per i termini di secondo grado.

Calcolando le derivate di P troviamo che

$$\nabla P(\mathbf{x}) = \mathbf{b} + 2C\mathbf{x}, \quad H_P(\mathbf{x}) = 2C. \quad (15)$$

5 Approssimazione locale di ordine 2

Grazie allo studio delle derivate seconde possiamo porci il problema dell'approssimazione (locale) di funzioni differenziabili con polinomi di secondo grado.

Definizione 5.1. Siano f e g due funzioni di classe C^2 definite in un intorno del punto $\mathbf{p} \in \mathbb{R}^n$. Diciamo che g approssima localmente f all'ordine 2 in un intorno di \mathbf{p} quando

$$f(\mathbf{x}) = g(\mathbf{x}) + o(\|\mathbf{x} - \mathbf{p}\|^2), \quad \text{per } \mathbf{x} \rightarrow \mathbf{p},$$

ovvero quando vale il limite

$$\lim_{\mathbf{x} \rightarrow \mathbf{p}} \frac{f(\mathbf{x}) - g(\mathbf{x})}{\|\mathbf{x} - \mathbf{p}\|^2} = 0.$$

Teorema 5.2. Siano f e g due funzioni scalari di classe C^2 definite in un intorno del punto $\mathbf{p} \in \mathbb{R}^n$. Allora sono equivalenti le seguenti proposizioni:

1. la funzione g approssima localmente f all'ordine 2 in un intorno di \mathbf{p} ;
2. nel punto \mathbf{p} i valori delle due funzioni e delle loro derivate prime e seconde coincidono, ovvero

$$f(\mathbf{p}) = g(\mathbf{p}), \quad \nabla f(\mathbf{p}) = \nabla g(\mathbf{p}), \quad H_f(\mathbf{p}) = H_g(\mathbf{p}).$$

Dimostriamo prima il caso particolare in cui g è identicamente nulla.

Lemma 5.3. Siano h una funzione scalare di classe C^2 definita in un intorno del punto $\mathbf{p} \in \mathbb{R}^n$. Allora sono equivalenti le seguenti proposizioni:

1. la funzione h è un infinitesimo di ordine 2 per $\mathbf{x} \rightarrow \mathbf{p}$, ovvero $h(\mathbf{x}) = o(\|\mathbf{x} - \mathbf{p}\|^2)$ per $\mathbf{x} \rightarrow \mathbf{p}$;
2. nel punto \mathbf{p} i valori della funzione e delle sue derivate prime e seconde sono nulli, ovvero

$$h(\mathbf{p}) = 0, \quad \nabla h(\mathbf{p}) = \mathbf{0}, \quad H_h(\mathbf{p}) = \mathbf{0}.$$

Dimostrazione. Sia \mathbf{u} un qualsiasi vettore non nullo di \mathbb{R}^n . Consideriamo la funzione

$$\varphi(t) := h(\mathbf{p} + t\mathbf{u}),$$

definita per t in un intorno di 0. Calcoliamo la derivata prima di φ , e utilizzando il fatto che h è differenziabile nel punto $\mathbf{p} + t\mathbf{u}$ otteniamo

$$\begin{aligned} \varphi'(t) &= \lim_{s \rightarrow 0} \frac{\varphi(t+s) - \varphi(t)}{s} = \lim_{s \rightarrow 0} \frac{h(\mathbf{p} + t\mathbf{u} + s\mathbf{u}) - h(\mathbf{p} + t\mathbf{u})}{s} = \\ &= D_{\mathbf{u}}h(\mathbf{p} + t\mathbf{u}) = \nabla h(\mathbf{p} + t\mathbf{u}) \cdot \mathbf{u} = \sum_{j=1}^n u_j \partial_{x_j} h(\mathbf{p} + t\mathbf{u}). \end{aligned}$$

Derivando ulteriormente, e utilizzando il fatto che anche le derivate parziali prime di f sono differenziabili nel punto $\mathbf{p} + t\mathbf{u}$ facendo calcoli analoghi ai precedenti, otteniamo

$$\varphi''(t) = \sum_{j=1}^n u_j \frac{d}{dt} (\partial_{x_j} h(\mathbf{p} + t\mathbf{u})) = \sum_{j=1}^n u_j \left(\sum_{k=1}^n u_k \partial_{x_k} \partial_{x_j} h(\mathbf{p} + t\mathbf{u}) \right) = \mathbf{u} \cdot (H_h(\mathbf{p} + t\mathbf{u})\mathbf{u}).$$

Quando $t = 0$, ricaviamo

$$\varphi(0) = h(\mathbf{p}), \quad \varphi'(0) = \nabla h(\mathbf{p}) \cdot \mathbf{u}, \quad \varphi''(0) = \mathbf{u} \cdot (H_h(\mathbf{p})\mathbf{u}). \quad (16)$$

Se prendiamo come ipotesi la proposizione 1 significa che vale il limite

$$\lim_{\mathbf{x} \rightarrow \mathbf{p}} \frac{h(\mathbf{x})}{\|\mathbf{x} - \mathbf{p}\|^2} = 0. \quad (17)$$

Restringiamo questo limite ai punti della retta passante per \mathbf{p} nella direzione \mathbf{u} , parametrizzata da $\mathbf{x} = \mathbf{p} + t\mathbf{u}$, e per $t \rightarrow 0$ otteniamo

$$\lim_{t \rightarrow 0} \frac{\varphi(t)}{t^2 \|\mathbf{u}\|^2} = \lim_{t \rightarrow 0} \frac{h(\mathbf{p} + t\mathbf{u})}{\|t\mathbf{u}\|^2} = 0.$$

E dunque $\varphi(t) = o(t^2)$ per $t \rightarrow 0$. Ma questo significa che il polinomio di McLaurin di ordine 2 per la funzione φ deve essere il polinomio nullo, e se i suoi coefficienti sono nulli ne segue che anche $\varphi(0)$, $\varphi'(0)$, $\varphi''(0)$ devono essere nulli, e per le formule (16) otteniamo che

$$h(\mathbf{p}) = 0, \quad \nabla h(\mathbf{p}) \cdot \mathbf{u} = 0, \quad \mathbf{u} \cdot H_h(\mathbf{p})\mathbf{u} = 0, \quad \forall \mathbf{u} \in \mathbb{R}^n.$$

In particolare se scegliamo $\mathbf{u} = \nabla h(\mathbf{p})$ otteniamo che

$$0 = \nabla h(\mathbf{p}) \cdot \nabla h(\mathbf{p}) = \|\nabla h(\mathbf{p})\|^2,$$

e dunque, avendo norma nulla, deve essere $\nabla h(\mathbf{p}) = 0$. Se scegliamo \mathbf{u} uguale ad un autovettore della matrice Hessiana con autovalore λ , per cui $H_h(\mathbf{p})\mathbf{u} = \lambda\mathbf{u}$, avremo che

$$0 = \mathbf{u} \cdot (H_h(\mathbf{p})\mathbf{u}) = \mathbf{u} \cdot (\lambda\mathbf{u}) = \lambda \|\mathbf{u}\|^2,$$

siccome un autovettore è sempre diverso da zero, non può avere norma nulla e dunque dovrà essere $\lambda = 0$. Quindi $H_h(\mathbf{p})$ è una matrice simmetrica con tutti gli autovalori nulli, e questo significa che deve essere la matrice nulla, $H_h(\mathbf{p}) = \mathbf{0}$, in quanto ogni matrice simmetrica è simile alla matrice diagonale che ha come elementi sulla diagonale gli autovalori, e l'unica matrice simile alla matrice nulla è la matrice nulla. Ecco provata così la proposizione 2.

Viceversa, prendiamo come ipotesi la proposizione 2. Essendo la funzione di classe C^2 abbiamo che le derivate parziali seconde $\partial_{x_j} \partial_{x_k} h$ sono continue e per ipotesi si annullano nel punto \mathbf{p} . Dunque anche la funzione

$$M(\mathbf{x}) := \sum_{j,k=1}^n |\partial_{x_j} \partial_{x_k} h(\mathbf{x})|$$

è continua e si annulla in \mathbf{p} . Questo significa che per ogni $\varepsilon > 0$ esiste un $\delta > 0$ tale che quando $\|\mathbf{x} - \mathbf{p}\| < \delta$ si ha $M(\mathbf{x}) < \varepsilon$. Fissiamo ora un punto \mathbf{x} nell'intorno di \mathbf{p} di raggio δ e consideriamo la funzione $\varphi(t) := h(\mathbf{p} + t(\mathbf{x} - \mathbf{p}))$ (è la stessa di prima ma con $\mathbf{x} - \mathbf{p}$ al posto di \mathbf{u}); abbiamo visto che si tratta di una funzione derivabile due volte, e per le formule (16) segue dall'ipotesi che

$$\varphi(0) = 0, \quad \varphi'(0) = 0, \quad \varphi''(0) = 0.$$

Applicando due volte il teorema di Lagrange otteniamo:

$$\begin{aligned} h(\mathbf{x}) &= \varphi(1) && \text{per definizione di } \varphi \\ &= \varphi(1) - \varphi(0) && \text{perché } \varphi(0) = 0 \\ &= \varphi'(t_1) && \exists t_1 \in]0, 1[\text{ per il teorema di Lagrange applicato a } \varphi \\ &= \varphi'(t_1) - \varphi'(0) && \text{perché } \varphi'(0) = 0 \\ &= \varphi''(t_2) && \exists t_2 \in]0, t_1[\text{ per il teorema di Lagrange applicato a } \varphi' \\ &= (\mathbf{x} - \mathbf{p}) \cdot H_h(\mathbf{p} + t_2(\mathbf{x} - \mathbf{p}))(\mathbf{x} - \mathbf{p}) && \text{per la formula (16) per } \varphi'' \text{ con } \mathbf{u} = \mathbf{x} - \mathbf{p} \end{aligned}$$

Per il $\boldsymbol{\xi} := \mathbf{p} + t_2(\mathbf{x} - \mathbf{p})$ abbiamo $\|\boldsymbol{\xi} - \mathbf{p}\| = t_2 \|\mathbf{x} - \mathbf{p}\| < t_2 \delta < \delta$, e dunque

$$\begin{aligned} \left| \frac{h(\mathbf{x})}{\|\mathbf{x} - \mathbf{p}\|^2} \right| &= \left| \frac{(\mathbf{x} - \mathbf{p}) \cdot H_h(\boldsymbol{\xi})(\mathbf{x} - \mathbf{p})}{\|\mathbf{x} - \mathbf{p}\|^2} \right| = \left| \sum_{j,k} \frac{x_j - p_j}{\|\mathbf{x} - \mathbf{p}\|} (\partial_{x_j} \partial_{x_k} h(\boldsymbol{\xi})) \frac{x_k - p_k}{\|\mathbf{x} - \mathbf{p}\|} \right| \leq \\ &\leq \sum_{j,k} \frac{|x_j - p_j|}{\|\mathbf{x} - \mathbf{p}\|} \left| \partial_{x_j} \partial_{x_k} h(\boldsymbol{\xi}) \right| \frac{|x_k - p_k|}{\|\mathbf{x} - \mathbf{p}\|} \leq \left(\sum_{j,k} |\partial_{x_j} \partial_{x_k} h(\boldsymbol{\xi})| \right) = M(\boldsymbol{\xi}) < \varepsilon, \end{aligned}$$

e questo implica che vale il limite (17) e quindi anche la proposizione 1. \square

Dimostrazione. [Dimostrazione del teorema 5.2] Consideriamo la differenza $h := f - g$; è anch'essa una funzione di classe C^2 e abbiamo

$$h(\mathbf{x}) := f(\mathbf{x}) - g(\mathbf{x}), \quad \nabla h(\mathbf{x}) := \nabla f(\mathbf{x}) - \nabla g(\mathbf{x}), \quad H_h(\mathbf{x}) := H_f(\mathbf{x}) - H_g(\mathbf{x}).$$

Le proposizioni 1 e 2 per le funzioni f e g del teorema sono equivalenti alle proposizioni 1 e 2 del lemma (5.3) per la funzione h . Dunque il teorema segue dal lemma. \square

Grazie a questi risultati possiamo determinare il polinomio di Taylor di ordine 2 per funzioni scalari di più variabili.

Teorema 5.4. *Sia f una funzione scalare di classe C^2 definita intorno al punto \mathbf{p} . Allora esiste ed unico un polinomio $T(\mathbf{x})$ di grado minore o uguale a due che approssima $f(\mathbf{x})$ al secondo ordine intorno a \mathbf{p} , ovvero per il quale si ha*

$$f(\mathbf{x}) = T(\mathbf{x}) + o(\|\mathbf{x} - \mathbf{p}\|^2), \quad \text{per } \mathbf{x} \rightarrow \mathbf{p}, \quad (18)$$

ed esso si dice polinomio di Taylor di ordine 2 per la funzione f intorno al punto \mathbf{p} , ed è dato da

$$T(\mathbf{x}) = f(\mathbf{p}) + \nabla f(\mathbf{p}) \cdot (\mathbf{x} - \mathbf{p}) + \frac{1}{2} (\mathbf{x} - \mathbf{p}) \cdot H_f(\mathbf{p})(\mathbf{x} - \mathbf{p}). \quad (19)$$

Dimostrazione. Vediamo prima il caso in cui $\mathbf{p} = \mathbf{0}$. Consideriamo un generico polinomio di secondo grado

$$T(\mathbf{x}) = a + \mathbf{b} \cdot \mathbf{x} + \mathbf{x} \cdot (C\mathbf{x}),$$

con coefficienti dati da un valore scalare $a \in \mathbb{R}$, un vettore $\mathbf{b} \in \mathbb{R}^n$ e una matrice $n \times n$ simmetrica C . Per calcolare le derivate di T utilizziamo le formule (15) e troviamo che il gradiente è dato da $\nabla T(\mathbf{x}) = \mathbf{b} + 2C\mathbf{x}$, mentre l'Hessiano è $H_T(\mathbf{x}) = 2C$. Per il teorema 5.2 abbiamo che T approssima f all'ordine 2 intorno a $\mathbf{0}$ se e solo se

$$f(\mathbf{0}) = T(\mathbf{0}) = a, \quad \nabla f(\mathbf{0}) = \nabla T(\mathbf{0}) = \mathbf{b}, \quad H_f(\mathbf{0}) = H_T(\mathbf{0}) = 2C,$$

e dunque se e solo se T è il polinomio

$$T(\mathbf{x}) = f(\mathbf{0}) + \nabla f(\mathbf{0}) \cdot \mathbf{x} + \frac{1}{2} \mathbf{x} \cdot H_f(\mathbf{0}) \mathbf{x}.$$

Per trattare il caso $\mathbf{p} \neq \mathbf{0}$ basta traslare il risultato ottenuto per $\mathbf{p} = \mathbf{0}$. Definiamo

$$\tilde{f}(\mathbf{y}) := f(\mathbf{p} + \mathbf{y}),$$

per \mathbf{y} in un intorno di $\mathbf{0}$. Se f è di classe C^2 intorno a \mathbf{p} allora \tilde{f} è di classe C^2 intorno a $\mathbf{0}$ e abbiamo

$$\nabla \tilde{f}(\mathbf{y}) = \nabla f(\mathbf{p} + \mathbf{y}), \quad H_{\tilde{f}}(\mathbf{y}) = H_f(\mathbf{p} + \mathbf{y}).$$

Per quanto visto nel caso $\mathbf{p} = \mathbf{0}$ abbiamo che l'unico polinomio che approssima \tilde{f} all'ordine 2 intorno a $\mathbf{0}$ è il polinomio

$$\tilde{T}(\mathbf{y}) := \tilde{f}(\mathbf{0}) + \nabla \tilde{f}(\mathbf{0}) \cdot \mathbf{y} + \frac{1}{2} \mathbf{y} \cdot H_{\tilde{f}}(\mathbf{0}) \mathbf{y}.$$

e abbiamo $\tilde{f}(\mathbf{y}) = \tilde{T}(\mathbf{y}) + o(\|\mathbf{y}\|^2)$ per $\mathbf{y} \rightarrow \mathbf{0}$. Effettuiamo la traslazione inversa: se sostituiamo $\mathbf{y} = \mathbf{x} - \mathbf{p}$ e poniamo $T(\mathbf{x}) := \tilde{T}(\mathbf{x} - \mathbf{p})$ otteniamo il polinomio (19) e l'approssimazione (18) per $f(\mathbf{x}) = \tilde{f}(\mathbf{x} - \mathbf{p})$. \square

Esempio 5.5. Determiniamo l'approssimazione di ordine 2 intorno al punto $(3, 2)$ per la funzione

$$f(x, y) := x \log(x - y).$$

Calcoliamo le derivate parziali prime e seconde per f ,

$$\begin{aligned} f_x(x, y) &= \log(x - y) + \frac{x}{x - y}, & f_y(x, y) &= -\frac{x}{x - y}, \\ f_{xx}(x, y) &= \frac{x - 2y}{(x - y)^2}, & f_{xy}(x, y) = f_{yx}(x, y) &= -\frac{y}{(x - y)^2}, & f_{yy}(x, y) &= -\frac{x}{(x - y)^2}. \end{aligned}$$

Nel punto $(3, 2)$ abbiamo

$$f(3, 2) = 0, \quad \nabla f(3, 2) = \begin{pmatrix} 3 \\ -3 \end{pmatrix}, \quad H_f(3, 2) = \begin{pmatrix} -1 & -2 \\ -2 & -3 \end{pmatrix}.$$

Il polinomio di Taylor di ordine 2 per f intorno al punto $(3, 2)$ è dunque

$$\begin{aligned} T(x, y) &= 0 + \binom{3}{-3} \cdot \begin{pmatrix} x-3 \\ y-2 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} x-3 \\ y-2 \end{pmatrix} \cdot \begin{pmatrix} -1 & -2 \\ -2 & -3 \end{pmatrix} \begin{pmatrix} x-3 \\ y-2 \end{pmatrix} \\ &= 3(x-3) - 3(y-2) - \frac{1}{2}(x-3)^2 - 2(x-3)(y-2) - \frac{3}{2}(y-2)^2 \\ &= -\frac{51}{2} + 10x + 9y - \frac{1}{2}x^2 - 2xy - \frac{3}{2}y^2. \end{aligned}$$

Troviamo così l'approssimazione

$$x \log(x-y) = -\frac{51}{2} + 10x + 9y - \frac{1}{2}x^2 - 2xy - \frac{3}{2}y^2 + o((x-3)^2 + (y-2)^2), \quad \text{per } (x, y) \rightarrow (3, 2).$$

Esempio 5.6. Determiniamo l'approssimazione del secondo ordine per la norma in \mathbb{R}^3 intorno al punto $\vec{i} = (1, 0, 0)$. Sia $f(x, y, z) = \|(x, y, z)\| = \sqrt{x^2 + y^2 + z^2}$. Abbiamo già calcolato in esempi precedenti il gradiente e l'Hessiano della norma,

$$\nabla f(x, y, z) = \|(x, y, z)\|^{-1} (x, y, z), \quad H_f(x, y, z) = \|(x, y, z)\|^{-3} \begin{pmatrix} y^2 + z^2 & -xy & -xz \\ -xy & x^2 + z^2 & -yz \\ -xz & -yz & x^2 + y^2 \end{pmatrix}.$$

Per il punto $\vec{i} = (1, 0, 0)$ abbiamo

$$f(\vec{i}) = \|\vec{i}\| = 1, \quad \nabla f(\vec{i}) = \vec{i} = (1 \ 0 \ 0), \quad H_f(\vec{i}) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

L'approssimazione cercata è data da

$$\begin{aligned} \|(x, y, z)\| &= 1 + (x-1) + \frac{1}{2}y^2 + \frac{1}{2}z^2 + o((x-1)^2 + y^2 + z^2) \\ &= x + \frac{1}{2}y^2 + \frac{1}{2}z^2 + o((x-1)^2 + y^2 + z^2), \quad \text{per } (x, y, z) \rightarrow (1, 0, 0). \end{aligned}$$

Esercizio 5.7. Determina l'approssimazione locale del secondo ordine per le seguenti funzioni intorno ai punti indicati.

1. $f(x, y) = x^3 + 2x^2y - 3xy^2 - 4y^3$ intorno al punto $(1, 2)$;
 2. $f(x, y) = \arctan \frac{y}{x}$ intorno al punto $(2, 0)$;
 3. $f(x, y) = \log(x^2 + y^2)$ intorno al punto $(-1, 1)$;
 4. $f(x, y, z) = \sqrt{x + yz}$ intorno al punto $(1, 1, 0)$;
 5. $f(x, y, z) = x \cos(y)e^z$ intorno al punto $(2, 0, 0)$;
 6. $f(x, y, z, t) = xyzt$ intorno al punto $(1, 2, 3, 4)$.
-

Lo studio dell'approssimazione al secondo ordine per funzioni differenziabili verrà ripreso durante l'insegnamento di Analisi Matematica 2 del secondo anno di corso per la classificazione dei punti critici e per la ricerca dei punti estremali (massimi e minimi locali) di una funzione di più variabili.