Università degli Studi di Palermo
Facoltà di Economia
CdS Statistica per l'Analisi dei Dati

Appunti del corso di Matematica

26 - Funzioni di più Variabili – Limiti e Derivate

Anno Accademico 2013/2014

M. Tumminello e A. Consiglio

1. Introduzione

In questa parte, vedremo come estendere alcuni dei concetti appresi per le funzioni di una variabile, come limite e derivata, a funzioni reali di più variabili. Per esempio:

1.1. Definizione di intorno in $\mathbb{R}^n$.
Sia $\overrightarrow{c} \in \mathbb{R}^n$ e $r \in \mathbb{R}$, $r > 0$. Si dice *intorno sferico* di $\overrightarrow{c}$ e raggio r l'insieme

$$B_r(\overrightarrow{c}) = \{\overrightarrow{x} \in \mathbb{R}^n \,:\, ||\overrightarrow{x} - \overrightarrow{c}|| < r\}$$

Ad esempio, in $\mathbb{R}^2$, l'intorno di $\overrightarrow{c}$ è un cerchio di raggio r centrato in $\overrightarrow{c}$.

1.2. Definizione di punto di accumulazione in $\mathbb{R}^n$. Sia $A \subseteq \mathbb{R}^n$ un insieme e $\overrightarrow{c} \in \mathbb{R}^n$. Diremo che $\overrightarrow{c}$ è un punto di accumulazione per A se, comunque preso un intorno $B_r(\overrightarrow{c})$ di $\overrightarrow{c}$, l'intersezione (escluso il punto $\overrightarrow{c}$) tra $B_r(\overrightarrow{c})$ e l'insieme A è non vuota. In simboli:

$$\forall r > 0, B_r(\overrightarrow{c}) \cap A \setminus \overrightarrow{c} \neq \phi$$

In altre parole, $\overrightarrow{c}$ è di accumulazione per A se è sempre possibile trovare punti di A, distinti da $\overrightarrow{c}$, in arbitrari intorni di $\overrightarrow{c}$, o, ancora, se in ogni intorno di $\overrightarrow{c}$ cadono infiniti punti di A.

1.3. Definizione di Limite di Funzioni di più variabili. Sia $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ una funzione e $\overrightarrow{c} \in \mathbb{R}^n$ un punto di accumulazione per A. Diremo che

$$\lim_{\overrightarrow{x} \to \overrightarrow{c}} f(\overrightarrow{x}) = L \in \mathbb{R}$$

se, $\forall \epsilon > 0$, $\exists r > 0$, r dipendente da ϵ, tale che $\forall \overrightarrow{x} \in B_r(\overrightarrow{c}) \bigcap A \setminus \overrightarrow{c}$, allora,

$$|f(\overrightarrow{x}) - L| < \epsilon$$

Come si può notare la definizione di limite è identica a quella per funzioni di una variabile, ed è possibile dimostrare che tutte le proprietà del limite si possono estendere alle funzioni $f : \mathbb{R}^n \to \mathbb{R}$. NOTA BENE: quanto affermato non è vero, in generale, per le funzioni vettoriali $\overrightarrow{f} : \mathbb{R}^n \to \mathbb{R}^m$.

Nel caso di funzioni $f : \mathbb{R} \to \mathbb{R}$ si è visto che il limite esiste se sono uguali il limite destro e il limite sinistro. Questa definizione di esistenza del limite non può essere generalizzata facilmente al caso di funzioni a più variabili, in quanto le possibili direzioni in cui convergere al punto di accumulazione sono infinite. Possiamo, comunque, utilizzare questo approccio per dimostrare che un limite non esiste, mostrando che il limite differisce per due direzioni diverse.

Esempio 1.1
Determinare, se esiste, il seguente limite:

$$\lim_{(x,y)\to(0,0)} \frac{x\,y}{x^2+y^2}$$

Indichiamo con $g(x,y)$ la nostra funzione: $g(x,y) = \frac{x\,y}{x^2+y^2}$. Scegliamo due direzioni specifiche per convergere al punto $(0,0)$ e mostriamo che i limiti sono diversi nelle due direzioni. In particolare, scegliamo le direzioni $(t,0)$ e (t,t), rispettivamente, lungo l'asse delle x e lungo la bisettrice del piano coordinato XY. Si osservi che $\lim_{t\to0}(t,0) = (0,0)$. Quindi:

$$\lim_{(t,0)\to(0,0)} g(t,0) = \lim_{t\to0} \frac{t\,0}{t^2+0} = 0.$$

In maniera analoga, abbiamo che $\lim_{t\to0}(t,t) = (0,0)$. Quindi:

$$\lim_{(t,t)\to(0,0)} g(t,t) = \lim_{t\to0} \frac{t\,t}{t^2+t^2} = \lim_{t\to0} \frac{t^2}{2\,t^2} = \frac{1}{2}.$$

Il limite differisce nelle due direzioni e quindi non esiste il limite per $(x,y)\to(0,0)$ di $g(x,y)$.

1.4. Definizione di Continuità in $\mathbb{R}^n$. Sia $f:\mathbb{R}^n\to\mathbb{R}$ definita in un intorno di $B^r(\overrightarrow{c})$, con $r\in\mathbb{R}$, $r>0$. La funzione $f(\overrightarrow{x})$ è continua in $\overrightarrow{c}$ se

$$\lim_{\overrightarrow{x}\to\overrightarrow{c}} f(\overrightarrow{x}) = f(\overrightarrow{c}).$$

Anche per la continuità si possono estendere tutti i teoremi che caratterizzano le proprietà delle funzioni continue di una variabile.

Esercizi 1.1

(1) Determinare i seguenti limiti e discutere la continuità di ciascuna funzione nel punto in cui il limite è calcolato:

$$\lim_{(x,y,z)\to(-2,1,0)} x\,e^{y\,z}; \qquad \lim_{(x,y)\to(2\,\pi,4)} \sin\left(\frac{x}{y}\right);$$

$$\lim_{(x,y)\to(-1,2)} \frac{x+y}{x-y}; \qquad \lim_{(x,y)\to(1,1)} \frac{x\,y}{x^2+y^2};$$

(2) Determinare, se esiste, il limite delle seguenti funzioni:

$$\lim_{(x,y)\to(0,0)} \frac{x-y}{\sqrt{x}-\sqrt{y}}; \qquad \lim_{(x,y)\to(0,0)} \frac{1}{x^2\,y^2};$$

$$\lim_{(x,y,z)\to(0,0,0)} \frac{x\,y+y\,z+x\,z}{x^2+y^2+z^2}; \qquad \lim_{(x,y,z)\to(0,0,0)} \frac{x\,y+y\,z^2+x\,z^2}{x^2+y^2+z^2};$$

Per le funzioni scalari la derivata di $f(x)$ nel punto c è definita dal limite, se esiste, del rapporto incrementale, per x che tende a c:

$$f'(c) = \lim_{x\to c} \frac{f(x)-f(c)}{x-c} = \lim_{h\to 0} \frac{f(c+h)-f(c)}{h}$$

Come abbiamo visto, in $\mathbb{R}^n$ è possibile tendere a $\overrightarrow{c}$ da infinite direzioni. Quindi il computo della derivata, ovvero, della pendenza in un punto deve tenere conto della direzione con cui si converge a $\overrightarrow{c}$.

Vediamo, con un esempio grafico, cosa significa tendere a $\overrightarrow{c}$ lungo una data direzione. Sia $f:\mathbb{R}^2\to\mathbb{R}$ una funzione definita da

$$f(x_1,x_2) = x_1^2 + x_2^2$$

e $\overrightarrow{c} = (c_1,c_2) = (1,1)$. Siano inoltre due direzioni identificate dai versori $\overrightarrow{w} = (w_1,w_2) = (-0.5/\sqrt{1.69},\ 1.2/\sqrt{1.69})$ e $\overrightarrow{t} = (t_1,t_2) = (1/\sqrt{1.25}, -0.5/\sqrt{1.25})$. Nella figura riportata di a pagina seguente (Fig.1) sono indicati i punti di $\mathbb{R}^2$ lungo le direzioni $\overrightarrow{w}$ e $\overrightarrow{t}$ ottenuti rispettivamente da

$$(c_1+h\,w_1, c_2+h\,w_2) \qquad \text{e} \qquad (c_1+h\,t_1, c_2+h\,t_2),$$

con $h\in\mathbb{R}$. Nella figura sono anche mostrate le curve sulla superficie $f(x_1,x_2)$ date da $f(c_1+h\,w_1, c_2+h\,w_2)$ e $f(c_1+h\,t_1, c_2+h\,t_2)$. Come si può notare dalla figura, la pendenza in $\overrightarrow{c}$ è diversa a seconda che si scelga la direzione la direzione $\overrightarrow{w}$ o $\overrightarrow{t}$ per convergere a $\overrightarrow{c}$.
Possiamo facilmente calcolare queste due pendenze, in quanto la derivata lungo una direzione, presa singolarmente, è equivalente alla derivata di una funzione scalare. Per esempio lungo la direzione $\overrightarrow{w}$ si ha che:

$$\lim_{h\to 0} \frac{f(\overrightarrow{c}+h\,\overrightarrow{w})-f(\overrightarrow{c})}{h} = \lim_{h\to 0} \frac{f(c_1+h\,w_1, c_2+h\,w_2)-f(c_1,c_2)}{h} =$$

$$\lim_{h\to 0} \frac{\left(1-\frac{0.5}{\sqrt{1.69}}\,h\right)^2 + \left(1+\frac{1.2}{\sqrt{1.69}}\,h\right)^2 - 2}{h} = \frac{2\cdot 0.7}{\sqrt{1.69}}$$

Si lascia allo studente di determinare la derivata lungo la direzione $\overrightarrow{t}$.

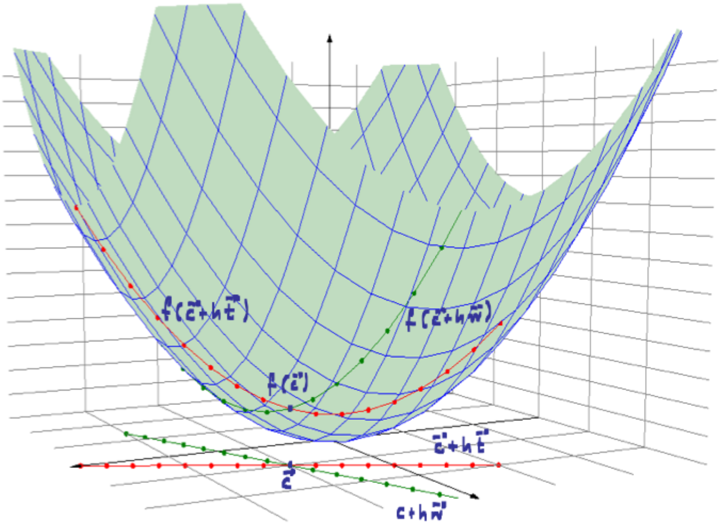

FIGURE 1. Rappresentazione delle funzioni $f(x_1, x_2) = x_1^2 + x_2^2$.

1.5. Definizione di Derivata Direzionale.

Sia $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ una funzione definita in un intorno aperto, $B_r(\overrightarrow{c})$, di $\overrightarrow{c} \in A \subseteq \mathbb{R}^n$ (punto di accumulazione per A). Sia, inoltre, $\overrightarrow{w} \in \mathbb{R}^n$ un versore. Si definisce derivata direzionale di $f(\overrightarrow{x})$ nel punto $\overrightarrow{c}$ e direzione $\overrightarrow{w}$ il limite per $h \to 0$ del rapporto incrementale lungo la direzione $\overrightarrow{w}$:

$$D_{\overrightarrow{w}} f(\overrightarrow{c}) = \lim_{h\to 0} \frac{f(\overrightarrow{c} + h\,\overrightarrow{w}) - f(\overrightarrow{c})}{h}$$

Di particolare interesse sono le derivate direzionali lungo le direzioni canoniche:

$$\overrightarrow{e}_k = (0, 0, \cdots, 0, 1, 0, \cdots, 0).$$

Si ricorda che il vettore $\overrightarrow{e}_k$ è un vettore con tutte le componenti nulle tranne la $k-$esima che è pari ad 1. Nel caso in cui $\overrightarrow{w} = \overrightarrow{e}_k$ la derivata direzionale diventa:

$$D_{\overrightarrow{e}_k} f(\overrightarrow{c}) = \lim_{h\to 0} \frac{f(\overrightarrow{c} + h\,\overrightarrow{e}_k) - f(\overrightarrow{c})}{h} =$$

$$= \lim_{h\to 0} \frac{f\left[(c_1, c_2, \cdots, c_k, \cdots, c_n) + h\,(0, 0, \cdots, 1, \cdots, 0)\right] - f(c_1, c_2, \cdots, c_k, \cdots, c_n)}{h} =$$

$$= \lim_{h\to 0} \frac{f\left[(c_1, c_2, \cdots, c_k + h, \cdots, c_n)\right] - f(c_1, c_2, \cdots, c_k, \cdots, c_n)}{h} =$$

$$= \lim_{h\to 0} \frac{g\,(c_k + h) - g(c_k)}{h} = g'(c_k).$$

Si osservi che la derivata lungo la $k-$esima direzione si ottiene considerando le rimanenti variabili come delle costanti. Infatti, la derivata direzionale $D_{\overrightarrow{e}_k} f(\overrightarrow{c})$ equivale alla derivata della funzione di una sola variabile $g(c_k)$.

1.6. Definizione di Derivata Parziale.
Sia $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ una funzione definita in un intorno aperto, $B_r(\overrightarrow{c})$, di $\overrightarrow{c} \in A \subseteq \mathbb{R}^n$ (punto di accumulazione per A). Sia, inoltre, $\overrightarrow{e}_k = (0, 0, \cdots, 1, \cdots, 0)$ il versore della $k-$esima direzione canonica, corrispondente alla $k-$esima variabile x_k. Si definisce *derivata direzionale rispetto ad* x_k, o *derivata parziale rispetto ad* x_k, nel punto $\overrightarrow{c}$, e si indica con

$$\frac{\partial f}{\partial x_k}(\overrightarrow{c}) \quad \text{oppure} \quad f_{x_k}(\overrightarrow{c}),$$

il limite, se esiste, di:

$$\lim_{h\to 0} \frac{f\left[(c_1, c_2, \cdots, c_k + h, \cdots, c_n)\right] - f(c_1, c_2, \cdots, c_k, \cdots, c_n)}{h}.$$

Da un punto di vista pratico, il calcolo della derivata parziale si ottiene derivando la funzione rispetto ad una variabile e considerando come costanti le rimanenti.

Esempi 1.2

- Determinare le derivate parziali della seguente funzione:
$$f(x_1, x_2, x_3) = 3\,x_1^2 - 4\,\sin(x_2) + 2\,x_1\,x_3.$$
Si ha che:
$$\frac{\partial f}{\partial x_1} = 6\,x_1 + 2\,x_3; \quad \frac{\partial f}{\partial x_2} = -4\,\cos(x_2); \quad \frac{\partial f}{\partial x_3} = 2\,x_1.$$
Nel punto $\overrightarrow{c} = (1, -1, 2)$, le derivate parziali assumeranno i seguenti valori:
$$\frac{\partial f}{\partial x_1}(1, -1, 2) = 10; \quad \frac{\partial f}{\partial x_2}(1, -1, 2) = -4\cdot\cos(-1); \quad \frac{\partial f}{\partial x_3}(1, -1, 2) = 2.$$
- Determinare le derivate parziali della seguente forma quadratica:
$$Q(x_1, x_2, x_3) = 3\,x_1^2 - 2\,x_1\,x_2 + x_1\,x_3 + x_2^2 - x_2\,x_3 + 4\,x_3^2.$$

Si ha:

$$\frac{\partial Q}{\partial x_1} = 6\,x_1 - 2\,x_2 + x_3; \quad \frac{\partial Q}{\partial x_2} = -2\,x_1 + 2\,x_2 - x_3; \quad \frac{\partial Q}{\partial x_3} = x_1 - x_2 + 8\,x_3.$$

Vediamo come questo risultato poteva essere ottenuto semplicemente dall'espressione matriciale della forma quadratica $Q(\overrightarrow{x})$. Abbiamo:

$$Q(\overrightarrow{x}) = \overrightarrow{x}^T A\,\overrightarrow{x} = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \cdot \begin{bmatrix} 3 & -1 & 1/2 \\ -1 & 1 & -1/2 \\ 1/2 & -1/2 & 4 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

D'altro canto, scrivendo le derivate parziali di cui sopra in notazione matriciale, abbiamo:

$$\frac{\partial Q}{\partial \overrightarrow{x}} = \begin{bmatrix} \frac{\partial Q}{\partial x_1} \\ \frac{\partial Q}{\partial x_2} \\ \frac{\partial Q}{\partial x_3} \end{bmatrix} = \begin{bmatrix} 6 & -2 & 1 \\ -2 & 2 & -1 \\ 1 & -1 & 8 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Confrontando questa espressione con l'espressione matriciale della forma quadratica data sopra, si vede che:

$$\frac{\partial Q}{\partial \overrightarrow{x}} = \frac{\partial(\overrightarrow{x}^T A\,\overrightarrow{x})}{\partial \overrightarrow{x}} = 2\,A\,\overrightarrow{x}.$$

QUESTO RISULTATO VALE, IN GENERALE, PER QUALSIASI FORMA QUADRATICA.

1.7. Definizione di funzione di classe C^1.

Sia $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ una funzione continua sull'insieme aperto A di $\mathbb{R}^n$. La funzione $f(\overrightarrow{x})$ è di classe C^1 su A, e si scrive $f \in C^1(A)$, se le derivate parziali $\frac{\partial f}{\partial x_k}$, $k = 1, \cdots, n$, sono continue su A.

1.8. Definizione di Gradiente.

Sia $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ una funzione definita in un intorno $B_r(\overrightarrow{c})$ del punto $\overrightarrow{c} \in A$. Esistano inoltre le derivate parziali $\frac{\partial f}{\partial x_k}(\overrightarrow{c})$, $k = 1, \cdots, n$ nel punto $\overrightarrow{c}$. Il vettore

$$\overrightarrow{\nabla} f(\overrightarrow{c}) = \left(\frac{\partial f}{\partial x_1}(\overrightarrow{c}), \frac{\partial f}{\partial x_2}(\overrightarrow{c}), \cdots, \frac{\partial f}{\partial x_n}(\overrightarrow{c}) \right)$$

è detto *gradiente* di f in $\overrightarrow{c}$.

TEOREMA 1.1. *Sia $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ una funzione di classe C^1 su A. Allora per ogni versore $\overrightarrow{w}$ la derivata direzionale di f nel punto $\overrightarrow{c}$ lungo la direzione $\overrightarrow{w}$ è data dal prodotto scalare tra il gradiente di f nel punto $\overrightarrow{c}$ e il vettore $\overrightarrow{w}$:*

$$D_{\overrightarrow{w}} f(\overrightarrow{c}) = \overrightarrow{\nabla} f(\overrightarrow{c}) \cdot \overrightarrow{w}.$$

Nell'esempio introduttivo, avevamo ottenuto che la derivata direzionale della funzione $f(x_1, x_2) = x_1^2 + x_2^2$, rispetto al vettore $\overrightarrow{w} = (w_1, w_2) = (-0.5/\sqrt{1.69},\ 1.2/\sqrt{1.69})$, calcolata nel punto $\overrightarrow{c} = (1, 1)$ era data da:

$$D_{\overrightarrow{w}} f(\overrightarrow{c}) = \frac{2 \cdot 0.7}{\sqrt{1.69}}.$$

Utilizzando il teorema precedente abbiamo che:

$$\frac{\partial f}{\partial x_1} = 2\, x_1; \quad \frac{\partial f}{\partial x_2} = 2\, x_2.$$

Quindi il gradiente di f nel punto $\overrightarrow{c} = (1, 1)$ sarà dato da:

$$\overrightarrow{\nabla} f(1, 1) = (2, 2).$$

La derivata direzionale sarà, secondo il teorema precedente, data da:

$$D_{\overrightarrow{w}} f(\overrightarrow{c}) = \overrightarrow{\nabla} f(\overrightarrow{c}) \cdot \overrightarrow{w} = \begin{bmatrix} 2 & 2 \end{bmatrix} \cdot \begin{bmatrix} -0.5/\sqrt{1.69} \\ 1.2/\sqrt{1.69} \end{bmatrix} = \frac{2 \cdot 0.7}{\sqrt{1.69}}$$

Questa proprietà della derivata direzionale permette di mettere in luce un'importante proprietà del vettore gradiente. In particolare, si ricorda che per qualsiasi coppia di vettori $\overrightarrow{u}$ e $\overrightarrow{v}$ si ha:

$$|\cos(\theta)| = \frac{|\overrightarrow{u} \cdot \overrightarrow{v}|}{||\overrightarrow{u}||\, ||\overrightarrow{v}||} \leq 1.$$

Applicando questa proprietà dei vettori alla coppia di vettori $\overrightarrow{\nabla} f(\overrightarrow{c})$ e $\overrightarrow{w}$ otteniamo che:

$$|D_{\overrightarrow{w}} f(\overrightarrow{c})| = |\overrightarrow{\nabla} f(\overrightarrow{c}) \cdot \overrightarrow{w}| = ||\overrightarrow{\nabla} f(\overrightarrow{c})||\, ||\overrightarrow{w}||\, |\cos(\theta)|.$$

Essendo $|\cos(\theta)| \leq 1$, il massimo valore di $|D_{\overrightarrow{w}} f(\overrightarrow{c})|$ si ottiene per $\theta = 0$ (oppure $\theta = \pi$), in altri termini, quando la direzione di $\overrightarrow{w}$ e del gradiente $\overrightarrow{\nabla} f(\overrightarrow{c})$ coincidono. Essendo $\overrightarrow{w}$ un versore ($||\overrightarrow{w}|| = 1$), ed essendo rilevante solo la sua direzione (non il verso), otterremo il mssimo valore di $|D_{\overrightarrow{w}} f(\overrightarrow{c})|$ quando:

$$\overrightarrow{w} = \frac{\overrightarrow{\nabla} f(\overrightarrow{c})}{||\overrightarrow{\nabla} f(\overrightarrow{c})||} \quad \text{oppure} \quad \overrightarrow{w} = -\frac{\overrightarrow{\nabla} f(\overrightarrow{c})}{||\overrightarrow{\nabla} f(\overrightarrow{c})||}.$$

Quindi, il gradiente di una funzione in un punto $\overrightarrow{c}$ identifica la *direzione* di massima variazione della funzione in quel punto e tale variazione risulterà pari alla norma del vettore gradiente:

$$|D_{\overrightarrow{w}} f(\overrightarrow{c})|_{MAX} = \left| \overrightarrow{\nabla} f(\overrightarrow{c}) \cdot \frac{\overrightarrow{\nabla} f(\overrightarrow{c})}{||\overrightarrow{\nabla} f(\overrightarrow{c})||} \right| = \left| \frac{||\overrightarrow{\nabla} f(\overrightarrow{c})||^2}{||\overrightarrow{\nabla} f(\overrightarrow{c})||} \right| = ||\overrightarrow{\nabla} f(\overrightarrow{c})||.$$

1.9. Derivabilità e approssimazione lineare di una funzione. Nella prima parte del corso abbiamo visto che, data una funzione $f : \mathbb{R} \to \mathbb{R}$, derivabile in un punto c, possiamo approssimare la variazione $\Delta f(c) = f(c+h) - f(c)$ utilizzando la derivata della funzione in c. Infatti, dalla definizione di derivata, si ha che:

$$\lim_{h \to 0} \frac{f(c+h) - f(c)}{h} = f'(c).$$

Se non effettuiamo il passaggio al limite, avremo che

$$f(c+h) - f(c) = f'(c)\,h + o(h),$$

dove il termine $o(h)$ è una quantità molto piccola rispetto ad h, più precisamente, una quantità che tende a zero per h che tende a zero più velocemente di h:

$$\lim_{h \to 0} \frac{o(h)}{h} = 0.$$

Questo ci assicura che, in un intorno di dimensione h di c la variazione di $f(x)$, df, è approssimata da una funzione lineare il cui coefficiente angolare è $f'(c)$. La quantità $df = f'(c)\,h$ è detta *differenziale* e possiamo approssimare $f(c+h) \simeq f(c) + f'(c)\,h$ per ogni valore piccolo di h. Dunque, se x è un punto "vicino" a c e ci chiediamo come approssimare la funzione f in x, noto il suo valore in c, f(c), avremo $f(x) \simeq f(c) + f'(c)\,(x-c) = A(x)$. La funzione $A(x) = f(c) + f'(c)\,(x-c) = f'(c)\,x + [f(c) - f'(c)\,c]$ è la somma di una *trasformazione lineare* $T(x) = f'(c)\,x$ e di un termine noto $[f(c) - f'(c)\,c]$.

Come si può generalizzare il concetto di differenziale per funzioni $f : \mathbb{R}^n \to \mathbb{R}^m$? In altri termini, come possiamo approssimare *df* tramite una funzione lineare in $\mathbb{R}^n$? Data l'analisi ora svolta per funzioni reali di una variabile reale il candidato ideale per approssimare una funzione $f : \mathbb{R}^n \to \mathbb{R}^m$ sarà una funzione $A(\overrightarrow{x})$ data dalla somma di una trasformazione lineare $T(\overrightarrow{x})$ e di un vettore noto $\overrightarrow{b}$: $A(\overrightarrow{x}) = T(\overrightarrow{x}) + \overrightarrow{b}$. Una funzione $A(\overrightarrow{x}) = T(\overrightarrow{x}) + \overrightarrow{b}$ che trasforma vettori di $\mathbb{R}^n$ in vettori di $\mathbb{R}^m$, dove $T(\overrightarrow{x})$ è una trasformazione lineare e $\overrightarrow{b}$ un vettore di $\mathbb{R}^m$, prende il nome di *trasformazione affine*. Abbiamo imparato che una qualsiasi trasformazione lineare $T(\overrightarrow{x})$ da $\mathbb{R}^n$ in $\mathbb{R}^m$ può essere scritta come $T(\overrightarrow{x}) = A\,\overrightarrow{x}$ dove A è una matrice di dimensione $m \times n$. Se siamo interessati a funzioni scalari, ovvero a funzioni $f : \mathbb{R}^n \to \mathbb{R}$, la matrice A si ridurrà, ponendo $m = 1$, ad un vettore $\overrightarrow{a}$. Dunque, per approssimare un funzione scalare $f : \mathbb{R}^n \to \mathbb{R}$ nell'intorno di un punto $\overrightarrow{c} \in \mathbb{R}^n$ la trasformazione affine $A : \mathbb{R}^n \to \mathbb{R}$ che ci interessa sarà del tipo $A(x) = \overrightarrow{a} \cdot \overrightarrow{x} + b$, dove b è uno scalare.

1.10. Definizione di derivabilità per funzioni di più variabili. Sia $f : \mathbb{R}^n \to \mathbb{R}$ definita in un intorno aperto del punto $\overrightarrow{c}$, $B_r(\overrightarrow{c})$. Se la funzione $f(\cdot)$ può essere approssimata nell'intorno del punto $\overrightarrow{c}$ tramite una trasformazione affine, diremo che $f(\overrightarrow{x})$ è *derivabile* nel punto $\overrightarrow{c}$. L'approssimazione tramite una trasformazione affine della funzione $f(\overrightarrow{x})$ in $\overrightarrow{c}$ sarà data da:

$$A(\overrightarrow{x}) = \overrightarrow{a} \cdot (\overrightarrow{x} - \overrightarrow{c}) + f(\overrightarrow{c}).$$

Si noti che questa trasformazione affine è del tipo che abbiamo considerato sopra: $A(x) = \overrightarrow{a} \cdot \overrightarrow{x} + b$. E' evidente che, per funzioni $f : \mathbb{R} \to \mathbb{R}$, il vettore $\overrightarrow{a}$ si riduce ad uno scalare a e risulta $a = f'(c)$. Questo vuol dire che, per funzioni di una variabile, se esiste la derivata prima in c allora la funzione è derivabile. In $\mathbb{R}^n$ la situazione non è altrettanto semplice. Ci chiediamo, innanzitutto: qual'è il vettore $\overrightarrow{a}$ tale che la trasformazione affine $A(\overrightarrow{x})$ meglio approssimi il valore di $f(\overrightarrow{x})$ nell'intorno del punto $\overrightarrow{c}$?

TEOREMA 1.2. *Sia $f : \mathbb{R}^n \to \mathbb{R}$ una funzione* derivabile *in $\overrightarrow{c}$. Allora la trasformazione affine che meglio approssima $f(\overrightarrow{x})$ nell'intorno del punto $\overrightarrow{c}$ è $A(\overrightarrow{x}) = \overrightarrow{a} \cdot (\overrightarrow{x} - \overrightarrow{c}) + f(\overrightarrow{c})$ con*

$$\overrightarrow{a} = \overrightarrow{\nabla} f(c).$$

Quindi la trasformazione affine che meglio approssima la funzione $f(\overrightarrow{x})$ nel intorno del punto $\overrightarrow{c}$ è

$$A(\overrightarrow{x}) = \overrightarrow{\nabla} f(c) \cdot (\overrightarrow{x} - \overrightarrow{c}) + f(\overrightarrow{c})$$

Vale la seguente condizione NECESSARIA affinché una funzione sia derivabile:

TEOREMA 1.3. *Sia $f : \mathbb{R}^n \to \mathbb{R}$ una funzione derivabile in $\overrightarrow{c}$. Allora $f(\overrightarrow{x})$ è continua in $\overrightarrow{c}$.*

Questo teorema indica che la continuità è una condizione necessaria (NON sufficiente) alla derivabilità.

Mentre per le funzioni scalari di variabile scalare l'esistenza della derivata prima implica che la funzione è derivabile nel punto in esame, nel caso di funzioni a più variabili l'esistenza delle derivate parziali NON è condizione sufficiente affinché la funzione $f(\overrightarrow{x})$ sia derivabile.

Esempio 1.3
Sia

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

Abbiamo già visto che $g(x,y)$ non ammette limite per $(x,y) \to (0,0)$. Quindi sicuramente $g(x,y)$ non è continua in $(0,0)$. Allora, per il teorema precedente, $g(x,y)$ non è derivabile in $(0,0)$. Graficamente (Fig.2), $g(x,y)$ ha un "taglio" in $(0,0)$, quindi non è possibile approssimare la funzione tramite un piano in $(0,0)$. Si osservi che, in questo caso le derivate parziali esistono ma non sono continue in $(0,0)$:

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

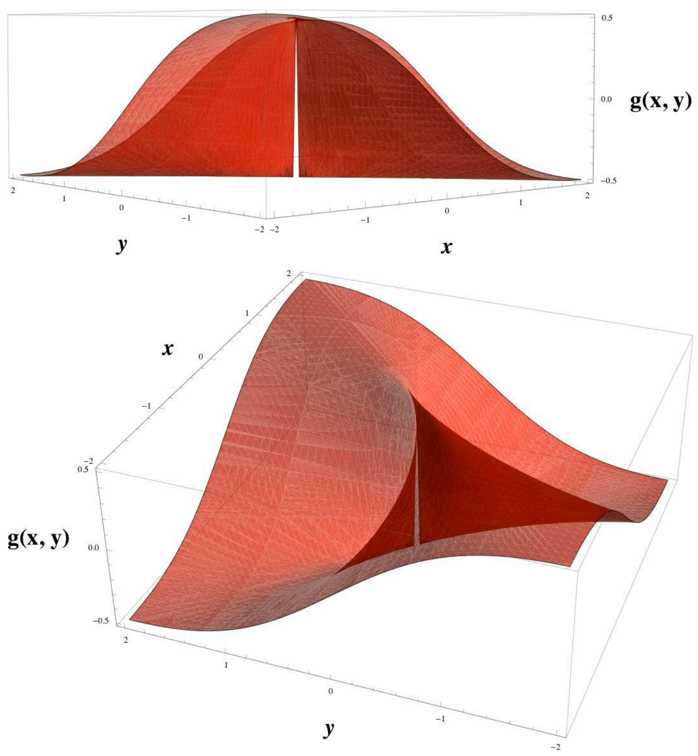

FIGURE 2. Rappresentazione, da due prospettive diverse, della funzione $g(x,y)$.

TEOREMA 1.4. *Sia $f : \mathbb{R}^n \to \mathbb{R}$ una funzione definita in un intorno del punto $\overrightarrow{c}$, $B_r(\overrightarrow{c})$. Sia inoltre $f(\overrightarrow{x})$ di classe $C^1(B_r(\overrightarrow{c}))$. Allora $f(\overrightarrow{x})$ è derivabile in $\overrightarrow{c}$.*

1.11. Definizione di Differenziale.
Sia $f : \mathbb{R}^n \to \mathbb{R}$ una funzione derivabile nel punto $\overrightarrow{c}$. Si definisce DIFFERENZIALE il termine:

$$df = \overrightarrow{\nabla} f(\overrightarrow{c}) \cdot \overrightarrow{h},$$

dove $\overrightarrow{c} + \overrightarrow{h} \in B_r(\overrightarrow{c})$. La funzione $f(\overrightarrow{c} + \overrightarrow{h})$ può essere approssimata come

$$f(\overrightarrow{c} + \overrightarrow{h}) \simeq f(\overrightarrow{c}) + \overrightarrow{\nabla} f(\overrightarrow{c}) \cdot \overrightarrow{h},$$

Nel caso di funzioni $f : \mathbb{R}^2 \to \mathbb{R}$, l'approssimazione della funzione nel punto $\overrightarrow{c}$ non è altro che il piano tangente alla funzione in $\overrightarrow{c}$.

Esempio 1.4
Data la funzione

$$f(x, y) = \frac{1}{2\,\pi}\, e^{-\frac{x^2+y^2}{2}},$$

determinare l'equazione del piano tangente alla funzione nel punto $\overrightarrow{c} = (0.5, 0.5)$ (vedi Fig.3 a pagina seguente).

Le derivate parziali sono:

$$\frac{\partial f}{\partial x} = \frac{-x}{2\,\pi}\, e^{-\frac{x^2+y^2}{2}}; \quad \frac{\partial f}{\partial y} = \frac{-y}{2\,\pi}\, e^{-\frac{x^2+y^2}{2}};$$

In $\overrightarrow{c}$ il gradiente è dunque dato da:

$$\overrightarrow{\nabla} f(0.5, 0.5) = \left(\left.\frac{\partial f}{\partial x}\right|_{(x,y)=(0.5,0.5)}, \left.\frac{\partial f}{\partial y}\right|_{(x,y)=(0.5,0.5)} \right) =$$

$$=(-0.06197, -0.06197),$$

e il piano tangente si ottiene da

$$\begin{aligned} z =& A(x, y) = f(0.5, 0.5) + \overrightarrow{\nabla} f(0.5, 0.5) \cdot (x - 0.5, y - 0.5) = \\ =& f(0.5, 0.5) + (-0.06197, -0.06197) \cdot (x - 0.5, y - 0.5) = \\ =& 0.1240 - 0.06197\,(x - 0.5) - 0.06197\,(y - 0.5) = \\ =& -0.06197\,x - 0.06197\,y + 0.1855. \end{aligned}$$

Come visto per le funzioni di una variabile reale, le derivate di ordine superiore al primo si ottengono sequenzialmente le derivate di ordine precedente. Così la derivata seconda è la derivata prima di $f'(x)$ e, in generale, $f^{(n)}(x) = \frac{d}{dx} f^{(n-1)}(x)$.

Nel caso di funzioni di più variabili, la derivazione può effettuarsi anche rispetto ad un'altra variabile. Per esempio, se $f : \mathbb{R}^2 \to \mathbb{R}$, le derivate parziali prime sono f_{x_1} e f_{x_2}. Possiamo derivare entrambe rispetto rispetto sia ad x_1 che ad x_2, ottenendo $f_{x_1\,x_1}$ e $f_{x_2\,x_2}$, ma anche $f_{x_1\,x_2}$ e $f_{x_2\,x_1}$. Le ultime due derivate sono dette *derivate parziali miste*. Altra possibile notazione per le derivate parziali di secondo ordine è la seguente:

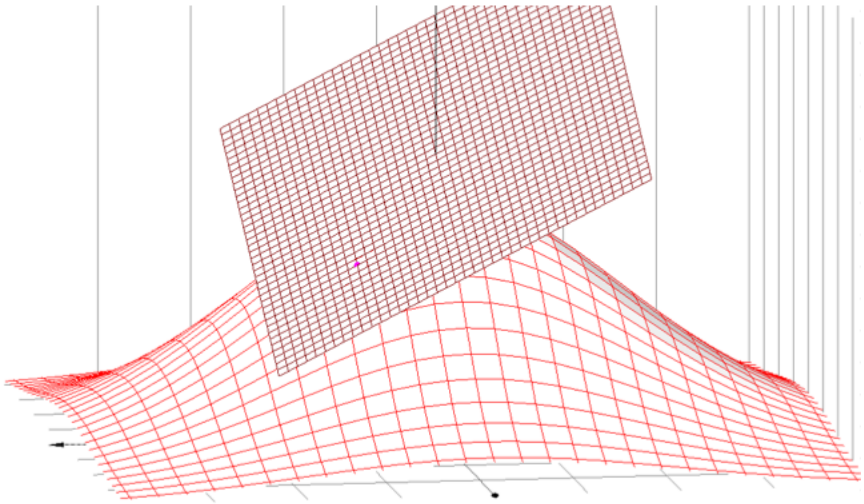

FIGURE 3. Rappresentazione della funzione $f(x, y)$ e del suo piano tangente nel punto $(0.5, 0.5)$.

$$\frac{\partial^2 f(x_1, x_2)}{\partial x_1^2} \quad \text{e per le derivate miste} \quad \frac{\partial^2 f(x_1, x_2)}{\partial x_1 \, \partial x_2}.$$

1.12. Definizione di Hessiano.
Sia $f : \mathbb{R}^n \to \mathbb{R}$ definita in un intorno $B_r(\overrightarrow{c})$ del punto $\overrightarrow{c} \in \mathbb{R}^n$. Se esistono tutte le derivate parziali di secondo ordine nel punto $\overrightarrow{c}$ si definisce **Hessiano** o **Matrice Hessiana** la matrice di dimensione $n \times n$ i cui elementi sono le derivate parziali $\frac{\partial^2 f(\overrightarrow{c})}{\partial x_i \, \partial x_j}$ per ogni $i, j = 1, \cdots, n$, e si indica con $D^2 f(\overrightarrow{c})$ oppure $H[f(\overrightarrow{c})]$, o, ancora come $H\, f(\overrightarrow{c})$:

$$H[f(\overrightarrow{c})] = \begin{bmatrix} \frac{\partial^2 f(\overrightarrow{c})}{\partial x_1^2} & \frac{\partial^2 f(\overrightarrow{c})}{\partial x_1 \, \partial x_2} & \cdots & \frac{\partial^2 f(\overrightarrow{c})}{\partial x_1 \, \partial x_n} \\ \frac{\partial^2 f(\overrightarrow{c})}{\partial x_2 \, \partial x_1} & \frac{\partial^2 f(\overrightarrow{c})}{\partial x_2^2} & \cdots & \frac{\partial^2 f(\overrightarrow{c})}{\partial x_2 \, \partial x_n} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial^2 f(\overrightarrow{c})}{\partial x_n \, \partial x_1} & \frac{\partial^2 f(\overrightarrow{c})}{\partial x_n \, \partial x_2} & \cdots & \frac{\partial^2 f(\overrightarrow{c})}{\partial x_n^2} \end{bmatrix}$$

1.13. Definizione di funzione di classe C^2.
Sia $f : \mathbb{R}^n \to \mathbb{R}$ una funzione continua sull'insieme aperto $A \subseteq \mathbb{R}^n$, e siano inoltre continue su A tutte le derivate parziali di secondo ordine $\frac{\partial^2 f}{\partial x_i \, \partial x_j}$ per ogni $i, j = 1, \cdots, n$. Allora diremo che $f(\overrightarrow{x})$ è di **classe** C^2 su A, e scriveremo $f \in C^2(A)$, oppure che $f(\overrightarrow{x})$ è di **tipo** C^2 su A.

Si noti che, in generale, $\frac{\partial^2 f}{\partial x_i \, \partial x_j} \neq \frac{\partial^2 f}{\partial x_j \, \partial x_i}$. Tuttavia, il seguente teorema mostra che per funzioni di classe C^2 risulta $\frac{\partial^2 f}{\partial x_i \, \partial x_j} = \frac{\partial^2 f}{\partial x_j \, \partial x_i}$.

TEOREMA 1.5. *Sia $f : \mathbb{R}^n \to \mathbb{R}$ una funzione di classe C^2 in un intorno del punto $\overrightarrow{c} \in \mathbb{R}^n$. Allora:*

$$\frac{\partial^2 f(\overrightarrow{c})}{\partial x_i \, \partial x_j} = \frac{\partial^2 f(\overrightarrow{c})}{\partial x_j \, \partial x_i}, \quad \forall i, j = 1, \cdots, n.$$