CALCOLO DIFFERENZIALE IN $\mathbb{R}^N$

1. *Limite per successioni in $\mathbb{R}^N$ e proprietà.*

Definizione 1.1. Sia $(P_n)_{n\in\mathbb{N}}$ una successione di punti in $\mathbb{R}^N$, ove, per ogni $n \in \mathbb{N}$, si ha $P_n = (x_1^n, \ldots, x_N^n)$. Si dice che la successione $(P_n)_{n\in\mathbb{N}}$ converge ad un punto $P_0 = (x_1^0, \ldots, x_N^0) \in \mathbb{R}^N$ se

$$\forall \varepsilon > 0 \ \ \exists\ n_0(\varepsilon) > 0 \ \ \text{tale che} \ \ d(P_n, P_0) = \|P_n - P_0\| < \varepsilon \ \ \forall n \geq n_0,$$

ove $d(P_n, P_0) = \|P_n - P_0\| = \sqrt{\sum_{i=1}^{N} (x_i^n - x_i^0)^2}$.

Si dice che la successione $(P_n)_{n\in\mathbb{N}}$ diverge a ∞, se $\|P_n\| \to +\infty$. Infine, si dice che la successione $(P_n)_{n\in\mathbb{N}}$ non è regolare se essa non è nè convergente nè divergente.

Anche per le successioni di punti di $\mathbb{R}^N$ valgono le seguenti proprietà : *unicità del limite*; *ogni successione convergente è limitata*; *criterio di Cauchy*; *Teorema di Bolzano–Weierstrass*; già viste nel caso delle successioni reali. Le corrispondenti dimostrazioni sono analoghe a quelle viste per le successioni reali, pur di sostituire al valore assoluto (cioè all'usuale distanza euclidea in $\mathbb{R}$) la corrispondente distanza euclidea di $\mathbb{R}^N$.

Teorema 1.2. *Sia $(P_n)_{n\in\mathbb{N}}$ una successione di punti in $\mathbb{R}^N$, ove, per ogni $n \in \mathbb{N}$, si ha $P_n = (x_1^n, \ldots, x_N^n)$ e sia $P_0 \in \mathbb{R}^N$. Allora $P_n \to P_0$ se e solo se $x_i^n \to x_i^0$, per ogni $i = 1, \ldots, N$.*

* * *

Dim. Dalla disuguaglianza (1.1) richiamata nelle dispense di topologia si ottiene

$$|x_i^n - x_i^0| \leq \sqrt{\sum_{i=1}^{N} (x_i^n - x_i^0)^2} \leq \sum_{i=1}^{N} |x_i^n - x_i^0|.$$

Pertanto, se $P_n \to P_0$, allora, per ogni $i = 1, \ldots, N$

$$|x_i^n - x_i^0| \leq \sqrt{\sum_{i=1}^{N} (x_i^n - x_i^0)^2} = \|P_n - P_0\| \to 0.$$

Viceversa, se $x_i^n \to x_i^0$, per ogni $i = 1, \ldots, N$, allora

$$\|P_n - P_0\| = \sqrt{\sum_{i=1}^{N} (x_i^n - x_i^0)^2} \leq \sum_{i=1}^{N} |x_i^n - x_i^0| \to 0.$$

Ciò conclude la dimostrazione. $\square$

* * *

2. Limite e continuità per funzioni di più variabili e proprietà.

Definizione 2.1. Sia A un sottoinsieme (aperto) di $I\!R^N$, sia $P_0 \in A$ un punto di accumulazione per A e sia $f : A \to I\!R$ una funzione assegnata. Diremo che $l \in I\!R$ è il limite di $f(P)$ per $P \to P_0$ se

$$\forall \varepsilon > 0 \ \ \exists\, \delta(\varepsilon) > 0 \ \text{ tale che } \ \forall P \in A \ \text{ con } \ 0 < d(P, P_0) = \|P - P_0\| < \delta$$
$$\text{allora } \ |f(P) - l| < \varepsilon.$$

In tal caso scriveremo

$$\lim_{P \to P_0} f(P) = l.$$

Analogamente, diremo che $+\infty$ (rispett. $-\infty$) è il limite di $f(P)$ per $P \to P_0$ e scriveremo

$$\lim_{P \to P_0} f(P) = +\infty \quad (\text{rispett. } -\infty)$$

se

$$\forall M > 0 \ \ \exists\, \delta(M) > 0 \ \text{ tale che } \ \forall P \in A \ \text{ con } \ 0 < d(P, P_0) = \|P - P_0\| < \delta$$
$$\text{allora } \ f(P) > M \ (\text{rispett. } < -M).$$

Anche per i limiti di funzioni di più variabili valgono le seguenti proprietà : *unicità del limite*; *Teoremi di permanenza del segno*; *Teorema ponte*; già viste nel caso delle successioni reali. Le corrispondenti dimostrazioni sono analoghe a quelle viste per le successioni reali, pur di sostituire (quando necessario) al valore assoluto (cioè all'usuale distanza euclidea in $I\!R$) la corrispondente distanza euclidea in $I\!R^N$.

Definizione 2.2. Sia A un sottoinsieme (aperto) di $I\!R^N$, sia $P_0 \in A$ e sia $f : A \to I\!R$ una funzione assegnata. Diremo che f è *continua* in P_0 se

$$\lim_{P \to P_0} f(P) = f(P_0)$$

ovvero se

$$\forall \varepsilon > 0 \ \ \exists\, \delta(\varepsilon) > 0 \ \text{ tale che } \ \forall P \in A \ \text{ con } \ d(P, P_0) = \|P - P_0\| < \delta$$
$$\text{allora } \ |f(P) - f(P_0)| < \varepsilon.$$

Definizione 2.3. Sia A un sottoinsieme (aperto) di $I\!R^N$ ed $f : A \to I\!R$ una funzione assegnata. Diremo che f è *uniformemente continua* in A se

$$\forall \varepsilon > 0 \ \ \exists\, \delta(\varepsilon) > 0 \ \text{ tale che } \ \forall P, Q \in A \ \text{ con } \ d(P, Q) = \|P - Q\| < \delta$$
$$\text{allora } \ |f(P) - f(Q)| < \varepsilon.$$

Anche per le funzioni continue di più variabili valgono le seguenti proprietà: *la composizione di funzioni continue è una funzione continua; se f, g sono due funzioni continue in P_0, allora le funzioni $f+g, f-g, fg$ sono continue in P_0; se f, g sono due funzioni continue in P_0 ed inoltre $g(P_0) \neq 0$, allora anche la funzione f/g è continua in P_0*; già viste nel caso delle funzioni di variabile reale.

Inoltre, per le funzioni continue definite su insiemi connessi valgono: il *Teorema degli zeri* ed i *Teoremi dei valori intermedi*; mentre per le funzioni continue definite su insiemi chiusi e limitati valgono il *Teorema di Weierstrass* ed il *Teorema di Heine–Cantor*.

Osservazione 2.4. Osserviamo che la continuità di f in un punto assegnato dipende dal comportamento della funzione nelle "immediate vicinanze" di tale punto.

Se consideriamo il caso delle funzioni reali $f : I \to I\!R$, ove I è un intervallo contenuto in $I\!R$, la continuità di f in un punto $x_0 \in I$ dipende dal comportamento di f nei punti x "vicini" ad x_0 e, poiché in questo caso x può muoversi solo lungo l'asse delle ascisse (esiste cioè una direzione privilegiata), è sufficiente studiare il comportamento di f a destra e a sinistra del punto x_0, stando "opportunamente vicino" ad esso.

Chiaramente la nozione di destra e sinistra perde senso quando siamo in $I\!R^N$ con $N > 1$; in questo caso non esiste più una direzione privilegiata e la continuità di f in un punto P_0 del suo dominio è determinata dal comportamento della funzione in tutto un intorno (opportunamente piccolo) del punto stesso. Questo rende in generale più difficile lo studio della continuità nel caso $N > 1$. Consideriamo, ad esempio, per $N = 2$ la funzione definita da

$$f(x,y) = \begin{cases} 0 & \text{se } xy = 0, \\ 1 & \text{altrimenti.} \end{cases} \tag{2.1}$$

In questo caso abbiamo che $\lim_{n\to+\infty} f(x_n, y_n) = f(0,0) = 0$ se $(x_n, y_n)_n$ è una successione infinitesima che "vive" sull'asse delle ascisse (cioè $(x_n, y_n) = (x_n, 0)$ con $x_n \to 0$) oppure su quello delle ordinate (cioè $(x_n, y_n) = (0, y_n)$ con $y_n \to 0$), ma la funzione NON è continua nell'origine, basta considerare il comportamento di f lungo una successione infinitesima che "vive", ad esempio, lungo la bisettrice del primo e terzo quadrante (cioè $x_n = y_n \to 0$). Una situazione ancora peggiore si ha con la funzione definita da

$$f(x,y) = \begin{cases} 1 & \text{se } y = x^2 \text{ e } (x,y) \neq (0,0) \\ 0 & \text{altrimenti.} \end{cases} \tag{2.2}$$

In questo caso, $\lim_{n\to+\infty} f(x_n, y_n) = f(0,0) = 0$ per ogni successione infinitesima $(x_n, y_n)_n$ che "vive" su una qualunque retta passante per l'origine, ma di nuovo la funzione NON è continua nell'origine, basta considerare il comportamento di f lungo una successione infinitesima $(x_n, y_n)_n$ tale che $y_n = x_n^2$.

Per ogni $i = 1, \ldots, N$, chiameremo i-esima proiezione $\pi_i : I\!R^N \to I\!R$, la funzione definita da $\pi_i(P) = x_i$, per ogni $P = (x_1, \ldots, x_N) \in I\!R^N$. Si può facilmente verificare (mediante la relazione (1.1) richiamata nelle dispense di topologia) che ogni proiezione i-esima è una funzione continua.

Definizione 2.5. Sia A un sottoinsieme (aperto) di $I\!R^N$ ed $f : A \to I\!R$ una funzione assegnata. Diciamo che f è *Lipschitziana* in A se esiste un numero reale $L > 0$ tale che per tutti i punti $P, Q \in A$ si abbia

$$|f(P) - f(Q)| \leq Ld(P, Q) = L\|P - Q\|.$$

Osservazione 2.6. Osserviamo innanzitutto che la definizione è del tutto analoga a quella vista nel caso delle funzioni di variabile reale; inoltre, anche in questo caso si può facilmente dimostrare che una funzione Lipschitziana è , in particolare, una funzione uniformemente continua in A e quindi anche continua in A.

3. Derivate parziali, gradiente, derivate direzionali.

Definizione 3.1. Sia $A \subseteq I\!R^N$ un insieme aperto, $P_0 = (x_1^0, \dots, x_N^0) \in A$ ed $f : A \to I\!R$ una funzione assegnata. Diremo che f è *derivabile parzialmente* rispetto alla variabile x_i nel punto $P_0 \in A$ se

$$\lim_{t\to 0} \frac{f(P_0 + t\mathrm{e}_i) - f(P_0)}{t} = \lim_{t\to 0} \frac{f(x_1^0, \dots, x_i^0 + t, \dots, x_N^0) - f(x_1^0, \dots, x_i^0, \dots, x_N^0)}{t}$$

esiste finito. In tal caso scriveremo

$$\frac{\partial f}{\partial x_i}(P_0) = \lim_{t\to 0} \frac{f(P_0 + t\mathrm{e}_i) - f(P_0)}{t}.$$

Osservazione 3.2. Se introduciamo la funzione di una variabile reale $\phi(t) = f(P_0 + t\mathrm{e}_i) = f(x_1^0, \dots, x_i^0 + t, \dots, x_N^0)$, si nota subito che la funzione f è parzialmente derivabile in P_0 rispetto ad x_i se e solo se la funzione ϕ è derivabile nel punto $t = 0$ ed inoltre

$$\phi'(0) = \frac{\partial f}{\partial x_i}(P_0).$$

In sostanza, la derivata parziale rispetto ad x_i di una assegnata funzione f in un punto P_0 è, se esiste, la derivata di f rispetto ad x_i, vista come unica variabile indipendente, mentre le altre variabili x_j, per $j \neq i$, sono pensate costanti.

Se una funzione $f : A \to I\!R$ ammette tutte le derivate parziali $\frac{\partial f}{\partial x_i}(P_0)$, per $i = 1, \dots, N$, in un punto P_0 dell'aperto $A \subseteq I\!R^N$, possiamo introdurre il vettore riga

$$\nabla f(P_0) = \Big(\frac{\partial f}{\partial x_1}(P_0), \dots, \frac{\partial f}{\partial x_N}(P_0)\Big)$$

detto *gradiente* della funzione f nel punto P_0.

ESEMPIO 3.3.

1) Sia $f : I\!R^2 \to I\!R$ definita da $f(x, y) = x^2 + 3xy$, allora

$$\frac{\partial f}{\partial x}(x, y) = 2x + 3y \qquad \frac{\partial f}{\partial y}(x, y) = 3x$$

da cui $\nabla f(x,y) = (2x+3y, 3x)$.

2) Sia $f : I\!R^3 \to I\!R$ definita da $f(x,y,z) = x\sin z + \cos(3xy)$, allora

$$\frac{\partial f}{\partial x}(x,y) = \sin z - 3y\sin(3xy) \qquad \frac{\partial f}{\partial y}(x,y) = -3x\sin(3xy) \qquad \frac{\partial f}{\partial z}(x,y) = x\cos z$$

da cui $\nabla f(x,y,z) = (\sin z - 3y\sin(3xy), -3x\sin(3xy), x\cos z)$.

3) Sia $A = I\!R \times (0,+\infty)$ ed $f : A \to I\!R$ definita da $f(x,y) = \log y$, allora

$$\frac{\partial f}{\partial x}(x,y) = 0 \qquad \frac{\partial f}{\partial y}(x,y) = \frac{1}{y}$$

da cui $\nabla f(x,y) = (0, \frac{1}{y})$.

Ovviamente se una funzione f è costante su un aperto $A \subseteq I\!R^N$, essa ammette tutte le derivate parziali ed esse risultano essere nulle.

Inoltre, per le funzioni di più variabili vale il seguente teorema, che generalizza il II Teorema della derivata nulla.

Teorema 3.4. *Sia $A \subseteq I\!R^N$ un insieme aperto e connesso. Supponiamo che $f : A \to I\!R$ sia una funzione che ammette tutte le derivate parziali in A e che tali derivate siano identicamente nulle in A. Allora f è costante in A.*

Abbiamo visto che per una funzione di più variabili reali si può introdurre, in un assegnato punto P_0 del suo dominio, la nozione di derivata parziale, che non è altro che il limite del rapporto incrementale lungo un asse coordinato. D'altra parte, gli assi coordinati individuano solo delle particolari direzioni passanti per P_0; in realtà per P_0 passano infinite direzioni, corrispondenti agli infiniti versori $v \in I\!R^N$. Quindi è ragionevole pensare di poter scrivere il rapporto incrementale lungo una generica direzione v e provare a farne il limite. In tal modo si introduce la nozione di *derivata direzionale*.

Definizione 3.5. Sia $A \subseteq I\!R^N$ un insieme aperto, $P_0 = (x_1^0, \ldots, x_N^0) \in A$, $v = (v_1, \ldots, v_N)$ un versore di $I\!R^N$ ed $f : A \to I\!R$ una funzione assegnata. Diremo che f è *derivabile lungo la direzione v* nel punto $P_0 \in A$ se

$$\lim_{t\to 0} \frac{f(P_0 + tv) - f(P_0)}{t} = \lim_{t\to 0} \frac{f(x_1^0 + tv_1, \ldots, x_N^0 + tv_N) - f(x_1^0, \ldots, x_N^0)}{t}$$

esiste finito. In tal caso scriveremo

$$\frac{\partial f}{\partial v}(P_0) = \lim_{t\to 0} \frac{f(P_0 + tv) - f(P_0)}{t}.$$

Ovviamente, se si sceglie come versore v quello relativo ad uno degli assi coordinati, cioè $v = \mathrm{e}_i$, la derivata direzionale lungo tale direzione coincide proprio con la derivata parziale rispetto ad x_i.

Infine osserviamo che per le funzioni di una variabile reale le nozioni di derivata parziale e derivata direzionale perdono senso, poichè si ha una sola direzione ammissibile, che è quella dell'asse x. Quindi l'unica definizione ragionevole è quella di derivata (senza ulteriori specificazioni), introdotta a suo tempo.

4. Differenziabilità e principali proprietà.

Abbiamo visto a suo tempo che, nel caso delle funzioni di una variabile reale, la nozione di derivabilità in un punto x_0 del dominio corrisponde all'esistenza della retta tangente al grafico della funzione in quel punto. Inoltre, una funzione derivabile in un punto risulta essere anche continua in tale punto. Quindi la derivabilità per le funzioni di variabile reale è una "buona" proprietà di regolarità.

Purtroppo, tutto ciò viene a mancare nel caso delle funzioni di più variabili reali; infatti in tal caso la derivabilità di una funzione in un punto P_0, intesa sia come derivabilità parziale che come derivabilità direzionale, eventualmente anche lungo tutte le possibili direzioni, non è una proprietà altrettanto buona. Più precisamente, è possibile costruire esempi di funzioni di più variabili che ammettono una o anche tutte le derivate parziali in un punto P_0, ma che non sono continue in P_0. Addirittura, ci sono esempi di funzioni di più variabili che ammettono tutte le derivate direzionali in un punto P_0, ma che non sono continue in P_0. A tale proposito, si considerino gli esempi (2.1) e (2.2), discussi precedentemente. Entrambe le funzioni non sono continue in $P_0 = (0,0)$, ma è facile verificare che, in entrambi i casi, le due derivate parziali esistono nell'origine e sono nulle. Addirittura, nell'esempio (2.2), si può verificare che nell'origine esistono anche tutte le derivate direzionali e sono nulle.

Per tali motivi, la nozione di derivabilità per funzioni di più variabili perde quel ruolo che aveva nel caso delle funzioni di una sola variabile. Pertanto siamo portati ad introdurre una nuova nozione (la *differenziabilità*), che risulterà essere effettivamente la nozione che estende al caso delle funzioni di più variabili il concetto di derivabilità delle funzioni di una variabile reale.

Definizione 4.1. Sia $A \subseteq \mathbb{R}^N$ un insieme aperto, $P_0 \in A$ ed $f : A \to \mathbb{R}$ una assegnata funzione. Diremo che f è *differenziabile* in P_0 se essa ammette tutte le derivate parziali in P_0 e vale la seguente relazione:

$$\lim_{\|H\|\to 0} \frac{f(P_0+H) - f(P_0) - \nabla f(P_0)\cdot H}{\|H\|} = 0 \qquad (H \in \mathbb{R}^N) \ .$$

Ricordando la definizione di "o" piccolo, si verifichi per esercizio che vale la seguente proposizione.

Proposizione 4.2. *Sia $A \subseteq \mathbb{R}^N$ un insieme aperto, $P_0 \in A$ ed $f : A \to \mathbb{R}$ una assegnata funzione. Allora f è differenziabile in P_0 se e solo se essa ammette tutte le derivate parziali in P_0 e vale la seguente relazione:*

$$f(P_0+H) = f(P_0) + \nabla f(P_0)\cdot H + o(\|H\|)$$

per ogni vettore $H \in \mathbb{R}^N$ di norma sufficientemente piccola in modo tale che $P_0 + H \in A$.

Poichè l'espressione $\nabla f(P_0)\cdot H$, fissato $P_0 \in A$, è una funzione lineare in H, quando f risulta essere differenziabile in P_0, chiameremo *piano tangente* al grafico di f nel punto P_0 il piano di equazione $z = f(P_0)+\nabla f(P_0)\cdot H$. La Proposizione 4.2 allora può essere interpretata geometricamente come segue: una funzione è differenziabile in un punto P_0 se e solo se il suo grafico ammette in tale punto un piano (detto, appunto, piano tangente) che approssimi la funzione nel punto in questione, a meno di infinitesimi di ordine superiore a $\|H\|$.

Infine, sottolineiamo che se f ammette tutte le derivate parziali in un punto P_0 (cioè ammette il vettore gradiente in P_0), è sempre possibile scrivere l'equazione del piano $z = f(P_0) + \nabla f(P_0)\cdot H$, però esso viene detto piano tangente SOLO se f risulta anche differenziabile in P_0, infatti in caso contrario tale piano NON approssima la funzione f in P_0 a meno di infinitesimi di ordine superiore a $\|H\|$ e quindi non è “degno” di tale nome.

Mostriamo ora che la nozione di differenziabilità è effettivamente l'estensione ad $I\!R^N$ della nozione di derivabilità in $I\!R$.

Teorema 4.3. *Sia $A \subseteq I\!R^N$ un insieme aperto, $P_0 = (x_1^0,\dots,x_N^0) \in A$ e sia $f : A \to I\!R$ una funzione assegnata. Assumiamo che f sia differenziabile in P_0. Allora valgono le seguenti proprietà:*

(i) f è continua in P_0;

(ii) f ammette tutte le derivate direzionali in P_0;

(iii) vale la seguente formula

$$\frac{\partial f}{\partial v}(P_0) = \nabla f(P_0)\cdot v = \sum_{i=1}^{N} \frac{\partial f}{\partial x_i}(P_0) v_i$$

per ogni versore $v = (v_1,\dots,v_N) \in I\!R^N$.

Dim.

(i) Dalla Proposizione 4.2 segue che, per $H \in I\!R^N$ sufficientemente piccolo in modo tale che $P_0 + H \in A$, si può scrivere

$$f(P_0+H) = f(P_0) + \nabla f(P_0)\cdot H + o(\|H\|). \tag{4.1}$$

Passando al limite per $\|H\| \to 0$ in entrambi i membri della precedente uguaglianza, si ottiene

$$\begin{aligned}\lim_{\|H\|\to 0} f(P_0+H) &= \lim_{\|H\|\to 0} [f(P_0) + \nabla f(P_0)\cdot H + o(\|H\|)] \\ &= f(P_0) + \lim_{\|H\|\to 0} [\nabla f(P_0)\cdot H + o(\|H\|)] = f(P_0)\end{aligned}$$

ove si è tenuto conto che

$$\lim_{\|H\|\to 0} \nabla f(P_0)\cdot H = 0 \qquad \text{e} \qquad \lim_{\|H\|\to 0} o(\|H\|) = 0.$$

Ciò dimostra che f è continua in P_0.

(ii) + (iii) Sia $v \in I\!R^N$ un qualunque versore e scegliamo $H = tv$ nella (4.1), con t sufficientemente piccolo in modo tale che $P_0 + tv \in A$. Allora

$$f(P_0 + tv) = f(P_0) + \nabla f(P_0) \cdot tv + o(\|tv\|).$$

e quindi

$$\frac{f(P_0 + tv) - f(P_0)}{t} = \frac{t\nabla f(P_0) \cdot v + o(|t|)}{t}.$$

Passando al limite per $t \to 0$ in entrambi i membri della precedente uguaglianza, si ottiene

$$\frac{\partial f}{\partial v}(P_0) = \lim_{t\to 0} \frac{f(P_0 + tv) - f(P_0)}{t} = \lim_{t\to 0} \frac{t\nabla f(P_0) \cdot v + o(|t|)}{t}$$
$$= \nabla f(P_0) \cdot v + \lim_{t\to 0} \frac{o(|t|)}{t} = \nabla f(P_0) \cdot v$$

che conclude la dimostrazione del teorema. □

Per quanto detto precedentemente, si osserva che l'esistenza di tutte le derivate parziali (ovvero del vettore gradiente) è una condizione necessaria per la differenziabilità , ma non sufficiente, in quanto essa non garantisce neppure la continuità. Vale però il seguente teorema, detto *Teorema del differenziale totale.*

Teorema 4.4. *Sia $A \subseteq I\!R^N$ un insieme aperto, $P_0 \in A$ ed $f : A \to I\!R$ una funzione assegnata. Assumiamo che f ammetta tutte le derivate parziali in un intorno di P_0 e che esse siano continue in P_0. Allora f è differenziabile in P_0.*

Se f ammette tutte le derivate parziali in A, continue in tutti i punti di A, allora si dice che f è di classe $\mathcal{C}^1(A)$. In particolare, dai precedenti teoremi si ottiene che, se $f \in \mathcal{C}^1(A)$, essa è differenziabile in tutto l'aperto A e quindi anche continua in A; pertanto, come nel caso delle funzioni di variabile reale, $\mathcal{C}^1(A) \subset \mathcal{C}(A)$.

5. Derivabilità delle funzioni composte.

Sia $(\phi_1(t), \ldots, \phi_N(t))$ una N-pla di funzioni definite su un intervallo reale I. Sia $A \subseteq I\!R^N$ un insieme aperto ed $f : A \to I\!R$ una funzione assegnata. Assumiamo che il codominio della N-pla di funzioni $(\phi_1, \ldots, \phi_N)$ sia contenuto in A, allora è possibile definire la funzione composta $h : I \to I\!R$ assegnata dall'espressione

$$h(t) = f(\phi_1(t), \ldots, \phi_N(t)).$$

Vale il seguente teorema.

Teorema 5.1. *Siano I, A, f, $\phi_1, \ldots, \phi_N$ ed h come sopra. Supponiamo che $P_0 \in A$ e $t_0 \in I$ siano tali che $(\phi_1(t_0), \ldots, \phi_N(t_0)) = P_0$. Allora*

(i) se f è continua in P_0 e $\phi_1, \ldots, \phi_N$ sono continue in t_0, anche la funzione h è continua in t_0;

(ii) se f è differenziabile in P_0 e $\phi_1, \ldots, \phi_N$ sono derivabili in t_0, anche la funzione h è derivabile in t_0 e si ha

$$h'(t_0) = \nabla f(P_0) \cdot (\phi'_1(t_0), \ldots, \phi'_N(t_0))^T = \sum_{i=1}^{N} \frac{\partial f}{\partial x_i}(P_0)\phi'_i(t_0),$$

ove $(\phi'_1(t_0), \ldots, \phi'_N(t_0))^T$ è il trasposto del vettore riga $(\phi'_1(t_0), \ldots, \phi'_N(t_0))$.

Osservazione 5.2. Sia $A \subseteq I\!R^N$ un insieme aperto ed $f : A \to I\!R$ una funzione differenziabile in $P_0 = (x_1^0, \ldots, x_N^0) \in A$. Sia $v \in I\!R^N$ un assegnato versore e definiamo per ogni $i = 1, \ldots, N$ le funzioni $\phi_i : I\!R \to I\!R$ date da

$$\phi_i(t) = x_i^0 + tv_i.$$

Ovviamente le funzioni ϕ_i, $i = 1, \ldots, N$, essendo lineari, sono continue e derivabili in $I\!R$, $(\phi_1(0), \ldots, \phi_N(0)) = P_0$ ed inoltre, se t varia in un intorno sufficientemente piccolo dell'origine, il codominio della N-pla $(\phi_1, \ldots, \phi_N)$ è contenuto in A. Allora possiamo definire la funzione composta $h(t) = f(\phi_1(t), \ldots, \phi_N(t))$ e dal precedente teorema (con $t_0 = 0$) si ottiene che h è derivabile nell'origine e vale la formula

$$h'(0) = \nabla f(P_0) \cdot (\phi_1'(0), \ldots, \phi_N'(0))^T = \sum_{i=1}^N \frac{\partial f}{\partial x_i}(P_0)\phi_i'(0) = \sum_{i=1}^N \frac{\partial f}{\partial x_i}(P_0) v_i. \tag{5.1}$$

Osservando che

$$\begin{aligned} h'(0) &= \lim_{t\to 0} \frac{h(t) - h(0)}{t} = \lim_{t\to 0} \frac{f(\phi_1(t), \ldots, \phi_N(t)) - f(\phi_1(0), \ldots, \phi_N(0))}{t} \\ &= \lim_{t\to 0} \frac{f(x_1^0 + tv_1, \ldots, x_N^0 + tv_N) - f(x_1^0, \ldots, x_N^0)}{t} \\ &= \lim_{t\to 0} \frac{f(P_0 + tv) - f(P_0)}{t} = \frac{\partial f}{\partial v}(P_0) \end{aligned} \tag{5.2}$$

da (5.1) e (5.2) si ritrova il risultato del Teorema 4.3.

Analogamente, sia ϕ una funzione definita su un intervallo reale I. Sia $A \subseteq I\!R^N$ un insieme aperto ed $f : A \to I\!R$ una funzione assegnata. Assumiamo che il codominio di f sia contenuto in I, allora è possibile definire la funzione composta $h : A \to I\!R$ assegnata dall'espressione

$$h(P) = \phi(f(P)).$$

Vale il seguente teorema.

Teorema 5.3. *Siano I, A, f, ϕ ed h come sopra. Supponiamo che $t_0 \in I$ e $P_0 \in A$ siano tali che $f(P_0) = t_0$. Allora*

(i) se ϕ è continua in t_0 ed f è continua in P_0, anche la funzione h è continua in P_0;

(ii) se ϕ è derivabile in t_0 ed f è differenziabile in P_0, anche la funzione h è differenziabile in P_0 e si ha

$$\frac{\partial h}{\partial x_i}(P_0) = \phi'(t_0) \frac{\partial f}{\partial x_i}(P_0) \qquad \forall i = 1, \ldots, N.$$

6. Derivate di ordine superiore al primo.

Sia $A \subseteq I\!R^N$ un insieme aperto ed $f : A \to I\!R$ una funzione che ammette tutte le derivate parziali in A. Se a loro volta le funzioni $f_{x_i} = \frac{\partial f}{\partial x_i} : A \to I\!R$, per $i = 1, \ldots, N$, ammettono derivate parziali

$$\frac{\partial f_{x_i}}{\partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}, \qquad i, j = 1, \ldots, N$$

queste ultime si dicono *derivate parziali seconde* della funzione f. In particolare le derivate parziali seconde $\frac{\partial^2 f}{\partial x_i \partial x_i} = \frac{\partial^2 f}{\partial x_i^2}$ si dicono *derivate seconde pure*, mentre le derivate parziali seconde $\frac{\partial^2 f}{\partial x_j \partial x_i}$ con $i \neq j$ si dicono *derivate seconde miste.*

In generale

$$\frac{\partial^2 f}{\partial x_j \partial x_i} \neq \frac{\partial^2 f}{\partial x_i \partial x_j} \qquad \text{per} \quad i \neq j$$

come per esempio si può verificare per la funzione

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

nell'origine. Vale però il seguente teorema, detto *Teorema di Schwarz.*

Teorema 6.1. *Sia $A \subseteq \mathbb{R}^N$ un insieme aperto, $P_0 \in A$ ed $f : A \to \mathbb{R}$ una funzione assegnata. Supponiamo che f ammetta le derivate seconde miste $\frac{\partial^2 f}{\partial x_i \partial x_j}$ e $\frac{\partial^2 f}{\partial x_j \partial x_i}$ in A e che esse siano continue in P_0, allora esse coincidono in P_0, cioè*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(P_0) = \frac{\partial^2 f}{\partial x_j \partial x_i}(P_0).$$

Se f ammette tutte le derivate parziali seconde in A, continue in tutti i punti di A, allora si dice che f è di classe $\mathcal{C}^2(A)$. In particolare, dai precedenti teoremi si ottiene facilmente che, se $f \in \mathcal{C}^2(A)$, allora $f \in \mathcal{C}^1(A)$.

Inoltre ogni funzione di classe $C^2(A)$ ha le derivate seconde miste uguali.

7. Studio degli estremanti di funzioni di più variabili.

Definizione 7.1. Sia $A \subseteq \mathbb{R}^N$, $P_0 \in A$ ed $f : A \to \mathbb{R}$ una funzione assegnata. Diremo che P_0 è punto di *massimo relativo* (rispett. *minimo relativo*) per la funzione f in A se esiste un intorno $B(P_0)$ del punto P_0 tale che

$$f(P) \leq f(P_0) \quad (\text{rispett. } f(P) \geq f(P_0)) \qquad \forall P \in B(P_0) \cap A\,. \tag{7.1}$$

Il punto P_0 si dice punto di *massimo assoluto* o *globale* (rispett. *minimo assoluto* o *globale*) se la disuguaglianza in (7.1) vale per tutti i punti $P \in A$ anzichè solo nell'insieme $B(P_0) \cap A$.

Se la disuguaglianza (7.1) è soddisfatta in senso stretto quando $P \neq P_0$, allora P_0 si dice punto di *massimo* (rispett. *minimo*) *relativo* o *assoluto stretto.* I punti di massimo e minimo relativo o assoluto sono anche detti punti *estremanti.*

Vediamo ora alcuni teoremi che ci saranno d'aiuto nella determinazione degli estremanti di una funzione di più variabili, qualora essa risulti abbastanza regolare.

Teorema 7.2. *Sia $A \subseteq \mathbb{R}^N$, $P_0 \in A$ ed $f : A \to \mathbb{R}$ una funzione dotata di tutte le derivate parziali in P_0. Allora, se P_0 è un estremante locale per f in A, il gradiente di f in P_0 è nullo.*

* * *

Dim. Supponiamo ad esempio che $P_0 = (x_1^0, \dots, x_N^0)$ sia un punto di minimo locale per f in A (il caso del massimo si fa in modo del tutto analogo). Allora, per ogni $i = 1, \dots, N$, la funzione di una variabile reale $h_i(t) = f(x_1^0, \dots, x_{i-1}^0, t, x_{i+1}^0, \dots, x_N^0)$ è definita in un intorno di x_i^0 ed ha un minimo locale proprio in $t_0 = x_i^0$. Ricordando che, se una funzione di una variabile reale ha un estremante in un punto t_0 interno al suo dominio ed in tale punto essa è derivabile, allora la sua derivata in t_0 è nulla, ne segue che $h_i'(x_i^0) = 0$ ovvero, ricordando la definizione appena data della funzione h_i, si ottiene

$$0 = h_i'(x_i^0) = \frac{\partial f}{\partial x_i}(P_0)\,.$$

Ciò conclude la dimostrazione del teorema. $\square$

* * *

Naturalmente, come per le funzioni di una sola variabile, la precedente condizione è solo una CONDIZIONE NECESSARIA per l'estremalità. Per esempio, la funzione $f : I\!R^2 \to I\!R$ definita da $f(x, y) = xy$ ha il gradiente nullo nell'origine, ma cambia segno in un intorno dell'origine (è positiva nel primo e terzo quadrante, mentre è negativa negli altri due quadranti), quindi il punto $(0, 0)$ non è un estremante per f in $I\!R^2$.

Definizione 7.3. Sia $A \subseteq I\!R^N$ un insieme aperto, $P_0 \in A$ ed $f : A \to I\!R$ una funzione dotata di gradiente in P_0. Se $\frac{\partial f}{\partial x_i}(P_0) = 0$ per ogni $i = 1, \dots, N$, diremo che P_0 è un *punto stazionario o critico* per f in A. Se f è una funzione differenziabile, i punti stazionari che non risultano essere estremanti si dicono *punti di sella* o semplicemente *selle.*

Da quanto detto precedentemente, segue che se f è una funzione differenziabile, ogni suo punto estremante è un punto stazionario, ma non vale il viceversa, come vedremo nei prossimi esempi.

Supponiamo ora che f sia addirittura di classe $\mathcal{C}^2(A)$. In tal caso possiamo introdurre la matrice quadrata $N \times N$ delle sue derivate seconde, detta *matrice Hessiana* di f, così definita:

$$Hf(P) = \begin{pmatrix} f_{x_1x_1}(P) & \dots & f_{x_1x_N}(P) \\ \dots & \dots & \dots \\ f_{x_Nx_1}(P) & \dots & f_{x_Nx_N}(P) \end{pmatrix}.$$

Dal Teorema di Schwarz, la matrice Hessiana di una funzione di classe $\mathcal{C}^2(A)$ risulta essere simmetrica, quindi, per un noto risultato di algebra lineare, essa ha tutti gli autovalori reali.

Ricordiamo inoltre che una matrice quadrata con tutti gli autovalori reali si dice *definita positiva* (rispett. *semidefinita positiva, definita negativa, semidefinita negativa*) se tutti i suoi autovalori sono strettamente positivi (rispett. maggiori o uguali a zero, strettamente negativi, minori o uguali a zero); essa si dice invece *indefinita* se possiede sia autovalori positivi che negativi.

Valgono i seguenti teoremi.

Teorema 7.4. *Sia $A \subseteq I\!R^N$ un insieme aperto, $P_0 \in A$ ed $f : A \to I\!R$ una funzione di classe $\mathcal{C}^2(A)$. Assumiamo che P_0 sia un punto stazionario per f in A. Allora si ha che*
(i) se $Hf(P_0)$ è definita positiva, allora P_0 è punto di minimo;
(ii) se $Hf(P_0)$ è definita negativa, allora P_0 è punto di massimo;
(iii) se $Hf(P_0)$ è indefinita, allora P_0 è punto di sella.

Teorema 7.5. *Sia $A \subseteq I\!R^N$ un insieme aperto, $P_0 \in A$ ed $f : A \to I\!R$ una funzione di classe $\mathcal{C}^2(A)$. Assumiamo che P_0 sia un punto stazionario per f in A. Allora si ha che*
(i) se P_0 è punto di minimo, allora $Hf(P_0)$ è semidefinita positiva;
(ii) se P_0 è punto di massimo, allora $Hf(P_0)$ è semidefinita negativa.

Osserviamo che il fatto che la matrice Hf sia definita positiva o negativa in un punto stazionario P_0 è SOLO un CONDIZIONE SUFFICIENTE per l'estremalità, ma NON è una CONDIZIONE NECESSARIA, come mostra il caso della funzione $f : I\!R^2 \to I\!R$ definita da $f(x,y) = x^4 + y^4$. Essa ha ovviamente un minimo assoluto nell'origine, ma nel punto $(0,0)$ si annullano sia il gradiente che tutti gli autovalori della matrice Hessiana. D'altra parte, il fatto che la matrice Hessiana sia solo semidefinita positiva o negativa in un punto stazionario, è una CONDIZIONE NECESSARIA ma NON SUFFICIENTE per l'estremalità; infatti tale punto potrebbe essere sia un estremante che una sella, come mostrano i seguenti esempi, per i quali l'origine è sempre un punto stazionario:
(1) $f(x,y) = x^3 + y^2$ ha un punto di sella nell'origine, dove la matrice Hessiana è semidefinita positiva;
(2) $f(x,y) = x^3 - y^2$ ha un punto di sella nell'origine, dove la matrice Hessiana è semidefinita negativa;
(3) $f(x,y) = x^4 + y^2$ ha un punto di minimo nell'origine, dove la matrice Hessiana è semidefinita positiva;
(2) $f(x,y) = -x^4 - y^2$ ha un punto di massimo nell'origine, dove la matrice Hessiana è semidefinita negativa.

Gli esempi precedenti mostrano che quando la matrice Hessiana è solo semidefinita, non è semplice studiare la natura del punto stazionario in questione, in tal caso bisognerà ricorrere ad altri metodi.