

Osservazioni sul differenziale

Luciano Battaia^(*)

11 ottobre 2010

Scopo di questo articolo è di presentare il concetto di differenziale per le funzioni reali di una o più variabili reali. L'intento non è quello di proporre una trattazione sistematica ed esaustiva, quanto piuttosto quello di evidenziare il significato e l'importanza del concetto di funzione differenziabile, in particolare per le funzioni di almeno due variabili.

Indice

Introduzione	1
1 Funzioni lineari	1
2 Funzioni di \mathbb{R} in \mathbb{R}	3
3 Funzioni di \mathbb{R}^2 in \mathbb{R}	7
4 Funzioni di \mathbb{R}^n in \mathbb{R}	9
5 La scrittura tradizionale	10
6 Funzioni di \mathbb{R}^n in \mathbb{R}^m	11
7 Algebra dei differenziali	12

Introduzione

Il concetto di differenziale è un concetto di straordinaria importanza sia teorica che pratica, in quanto solo le funzioni differenziabili sono dotate di tangente (o di piano tangente, o di iperpiano tangente), cioè solo le funzioni differenziabili sono “ben approssimabili”, in un intorno del punto dove si calcola il differenziale, con una funzione particolarmente semplice, come sono le funzioni polinomiali di primo grado.

L'importanza del differenziale emerge tuttavia solo con le funzioni (reali) di almeno due variabili, in quanto nel caso delle funzioni di una sola variabile reale una funzione è differenziabile se e solo se è derivabile, e siccome il concetto di derivata appare più semplice, di solito è l'unico ad essere trattato. Inoltre, a mio modo di vedere, le strategie usate per introdurre il concetto di differenziale sono spesso poco chiare e fuorvianti; se si usa la dovuta precisione terminologica e si chiariscono accuratamente tutti gli aspetti, il concetto di differenziale cesserà di essere quell'“oggetto misterioso” che è per la maggior parte degli studenti.

Per chiarire quanto sopra affermato riporto la definizione di differenziale (per funzioni reali di una variabile reale) presa da un testo molto in voga (si tratta comunque della definizione più diffusa): *Il*

^{*}<http://www.batmath.it>

differenziale di una funzione è il prodotto della derivata della funzione per l'incremento arbitrario della variabile indipendente. Dopo aver letto questa definizione (e magari aver anche esaminato alcuni esempi) credo sia abbastanza difficile rispondere a una domanda del tipo: ma il differenziale è un numero, oppure è una funzione? Se è una funzione, di quante variabili? E qual è il suo grafico cartesiano?

1 Funzioni lineari

Le funzioni lineari sono le funzioni fondamentali nello studio degli spazi vettoriali. Ne richiamiamo brevemente la definizione che ci servirà in quanto il differenziale di una funzione è una funzione lineare, limitandoci al caso di spazi vettoriali su \mathbb{R} .

Definizione 1. Dati due spazi vettoriali V e W , entrambi sul corpo dei numeri reali, una funzione $f: V \rightarrow W$ si dice lineare se

$$(1) \quad f(\lambda \vec{u} + \mu \vec{v}) = \lambda f(\vec{u}) + \mu f(\vec{v}), \quad \forall \lambda, \mu \in \mathbb{R} \text{ e } \forall \vec{u}, \vec{v} \in V.$$

La definizione data significa che se una funzione è lineare è indifferente scegliere di eseguire prima alcune operazioni (lineari!) su due vettori di V e poi applicare la funzione, oppure applicare prima la funzione e poi eseguire le corrispondenti operazioni sulle immagini dei due vettori di V .

Nel caso di funzioni di \mathbb{R}^n in \mathbb{R} le funzioni lineari sono sempre di uno dei tipi seguenti:

$$(2) \quad f: \mathbb{R}^n \rightarrow \mathbb{R}, \quad \begin{cases} f(x_1) & = & a_1 x_1 & (n=1) \\ f(x_1, x_2) & = & a_1 x_1 + a_2 x_2 & (n=2) \\ f(x_1, x_2, x_3) & = & a_1 x_1 + a_2 x_2 + a_3 x_3 & (n=3) \\ \dots & \dots & \dots & \dots \end{cases}.$$

Le funzioni di \mathbb{R}^n in \mathbb{R}^m possono essere sempre pensate come una m -upla di funzioni di \mathbb{R}^n in \mathbb{R} , e quindi le funzioni lineari saranno m -uple di funzioni di uno dei tipi sopra considerati:

$$(3) \quad f: \mathbb{R}^n \rightarrow \mathbb{R}^m, \quad \begin{cases} f_1(x_1, \dots, x_n) & = & a_{11}x_1 + \dots + a_{1n}x_n \\ f_2(x_1, \dots, x_n) & = & a_{21}x_1 + \dots + a_{2n}x_n \\ \vdots & \vdots & \vdots \\ f_m(x_1, \dots, x_n) & = & a_{m1}x_1 + \dots + a_{mn}x_n \end{cases}$$

Posto

$$(4) \quad \vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \text{e} \quad A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix},$$

è ben noto che la funzione si può scrivere, in forma matriciale, come

$$(5) \quad f: \mathbb{R}^n \rightarrow \mathbb{R}^m, \quad \vec{x} \mapsto A\vec{x}.$$

Per comodità di scrittura, comunque, scriveremo il vettore \vec{x} sempre come vettore riga anzichè come vettore colonna, $\vec{x} = (x_1, \dots, x_n)$.

Gli spazi vettoriali \mathbb{R}^n , per le applicazioni che ci interessano, sono pensati anche dotati del prodotto scalare canonico e della norma dedotta dal prodotto scalare.

Nel caso delle funzioni di \mathbb{R}^2 o di \mathbb{R}^3 in \mathbb{R} è tradizione indicare le variabili con (x, y) e (x, y, z) rispettivamente, anzichè con (x_1, x_2) oppure (x_1, x_2, x_3) . Poichè però in questo articolo vogliamo

proporre una trattazione unificata dei vari casi, manterremo la notazione con gli indici, riservando la y (ed eventualmente $\vec{y} = (y_1, \dots, y_m)$) per la variabile dipendente.

Per le funzioni a valori in \mathbb{R} , se si usa una rappresentazione cartesiana, le funzioni lineari hanno un grafico particolarmente semplice e significativo, come di seguito indicato.

- Una retta per l'origine del piano Ox_1y , nel caso $n = 1$, e il numero a_1 rappresenta il coefficiente angolare della retta; se poi si scrive l'equazione della retta in forma implicita (o normale), $a_1x_1 - y = 0$, si deduce che il vettore $(a_1, -1)$ è un vettore perpendicolare alla retta.
- Un piano per l'origine dello spazio Ox_1x_2y nel caso di $n = 2$; se si scrive l'equazione del piano in forma implicita (o normale), $a_1x_1 + a_2x_2 - y = 0$, si vede che il vettore $(a_1, a_2, -1)$ è un vettore perpendicolare al piano.
- Un iperpiano per l'origine dello spazio quadridimensionale $Ox_1x_2x_3y$ nel caso di $n = 3$; se si scrive l'equazione dell'iperpiano in forma implicita (o normale), $a_1x_1 + a_2x_2 + a_3x_3 - y = 0$, si vede che il vettore $(a_1, a_2, a_3, -1)$ è un vettore perpendicolare all'iperpiano.

2 Funzioni di \mathbb{R} in \mathbb{R}

Come già accennato, nel caso di funzioni reali di una sola variabile reale, il concetto di differenziabilità non introduce sostanziali novità rispetto a quello di derivabilità. Tuttavia è opportuno trattare ugualmente in dettaglio il problema, in quanto in questa situazione l'interpretazione grafico-geometrica è particolarmente semplice e significativa.

Sia $f: I \subset \mathbb{R} \rightarrow \mathbb{R}$ una funzione reale di una variabile reale; per semplicità supporremo che il dominio I sia un intervallo aperto, di modo che i punti che considereremo saranno sempre interni all'intervallo. Supporremo inoltre che la funzione f sia *continua* in I . Sia x_0 un punto di I , e $f(x_0)$ il corrispondente valore della funzione. Siamo interessati a valutare la variazione della funzione al passaggio da x_0 a un punto generico x situato nei pressi di x_0 ; se poniamo

$$(6) \quad h = x - x_0,$$

questa variazione sarà data da

$$(7) \quad \Delta f_{x_0}(h) = f(x_0 + h) - f(x_0),$$

che è una funzione di h , una volta fissato x_0 . Questa variazione prende il nome di *incremento* della funzione f relativo all'incremento h della variabile e al punto x_0 .

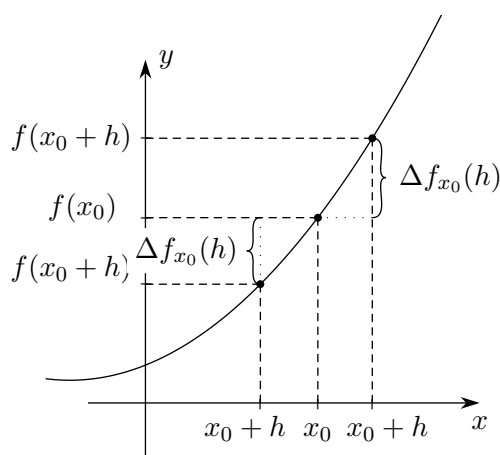


Figura 1 Incremento $\Delta f_{x_0}(h)$ di una funzione

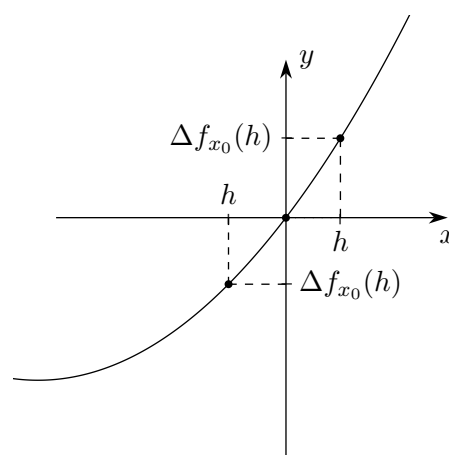


Figura 2 Incremento $\Delta f_{x_0}(h)$ in funzione di h

Nella figura 1 è evidenziato l'incremento a partire da un dato punto x_0 (questo incremento è positivo se $f(x_0 + h) > f(x_0)$, altrimenti è negativo, e il tutto dipende sia dal segno di h , cioè dal fatto che $x_0 + h$ sia a destra o a sinistra di x_0 sia dal tipo di funzione considerata); nella figura 2 è rappresentato lo stesso incremento in funzione di h : si tratta semplicemente della traslazione rigida del grafico della figura 1 ottenuta spostando il punto $(x_0, f(x_0))$ sull'origine.

La domanda che dà origine al concetto di differenziale è la seguente: è possibile approssimare l'incremento, cioè la funzione il cui grafico è rappresentato nella figura 2, con una funzione lineare, ovvero una funzione avente per grafico una retta passante per l'origine? Per rispondere alla domanda consideriamo una generica funzione lineare, $x \mapsto L_{x_0}(x)$ e valutiamo l'errore che si commette calcolando $L_{x_0}(h)$ al posto di $\Delta f_{x_0}(h)$, errore che è dato da

$$(8) \quad \omega(h) = \Delta f_{x_0}(h) - L_{x_0}(h).$$

Nella figura 3 che segue è visualizzata la situazione corrispondente a una particolare scelta della funzione L_{x_0} . Naturalmente $\omega(h)$ sarà positivo o negativo a seconda che $\Delta f_{x_0}(h)$ sia maggiore o minore di $L_{x_0}(h)$.

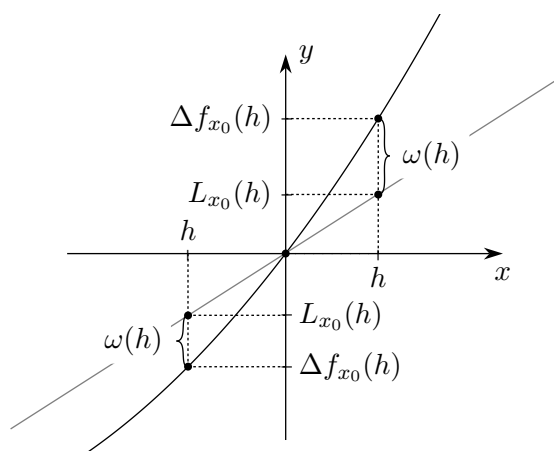


Figura 3 Una possibile funzione lineare, L_{x_0} , approssimante $\Delta f_{x_0}(h)$

Nelle ipotesi di continuità⁽¹⁾ in cui ci siamo posti, è chiaro che

$$(9) \quad \lim_{h \rightarrow 0} \omega(h) = \lim_{h \rightarrow 0} \Delta f_{x_0}(h) - L_{x_0}(h) = \lim_{h \rightarrow 0} (f(x_0 + h) - f(x_0)) - L_{x_0}(h) = 0,$$

ovvero che ω è una funzione infinitesima, al tendere di h a 0.

Ci possiamo ora chiedere se è possibile scegliere la funzione lineare L_{x_0} in modo che ω sia addirittura un *infinitesimo di ordine superiore* rispetto ad h , ovvero in modo tale che

$$(10) \quad \lim_{h \rightarrow 0} \frac{\omega(h)}{h} = \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0) - L_{x_0}(h)}{h} = \lim_{h \rightarrow 0} \frac{\Delta f_{x_0}(h) - L_{x_0}(h)}{h} = 0.$$

L'esempio seguente prova che, in genere, la risposta è negativa. Consideriamo la funzione

$$(11) \quad f: \mathbb{R} \rightarrow \mathbb{R}, \quad f(x) = \begin{cases} x \sin \frac{1}{x}, & \text{se } x \neq 0 \\ 0, & \text{se } x = 0 \end{cases},$$

¹La continuità serve per concludere che

$$\lim_{h \rightarrow 0} f(x_0 + h) = f(x_0).$$

e una generica funzione lineare, $L_{x_0}(h) = mh$. Sia inoltre $x_0 = 0$. Si ha allora

$$(12) \quad \omega(h) = \begin{cases} h \sin \frac{1}{h} - mh, & \text{se } h \neq 0 \\ 0, & \text{se } h = 0 \end{cases},$$

da cui

$$(13) \quad \lim_{h \rightarrow 0} \frac{\omega(h)}{h} = \lim_{h \rightarrow 0} \frac{h \sin \frac{1}{h} - mh}{h} = \lim_{h \rightarrow 0} \left(\sin \frac{1}{h} - m \right),$$

ed è ben noto che questo limite non esiste, qualunque sia m .

Se invece consideriamo la funzione data da $f(x) = x^2$, il punto $x_0 = 1$, e ancora una generica funzione lineare $L_{x_0}(h) = mh$, otteniamo

$$(14) \quad \omega(h) = ((1+h)^2 - 1) - mh = h^2 + 2h - mh,$$

da cui

$$(15) \quad \lim_{h \rightarrow 0} \frac{\omega(h)}{h} = \lim_{h \rightarrow 0} \frac{h^2 + 2h - mh}{h} = \lim_{h \rightarrow 0} (h + 2 - m),$$

e si conclude facilmente che questo limite è zero se, e solo se, $m = 2$. Poiché 2 è esattamente la derivata della funzione f in esame, in corrispondenza al punto $x_0 = 1$, ne concludiamo che

$$(16) \quad \omega(h) = \Delta f_1(h) - L_{x_0}(h) = f(1+h) - f(1) - f'(1)h \Rightarrow f(1+h) = f(1) + f'(1)h + \omega(h).$$

Se teniamo conto, vedi la formula (6) nella pagina 3, che $h = x - 1$, otteniamo

$$(17) \quad f(x) = f(1) + f'(1)(x-1) + \omega(x-1),$$

ovvero: *la funzione, nei pressi del punto 1, può essere approssimata dalla retta tangente al grafico, a meno di un infinitesimo di ordine superiore rispetto a $x - 1$.*

La situazione che si verifica per la funzione $f(x) = x^2$ è di grande importanza e porta alla introduzione del concetto di differenziale.

Definizione 2. Sia $f: I \rightarrow \mathbb{R}$ una funzione definita in un (intervallo) aperto I di \mathbb{R} . La funzione f si dice differenziabile in un punto $x_0 \in I$ se esiste una funzione lineare $L_{x_0}: \mathbb{R} \rightarrow \mathbb{R}$ tale che

$$(18) \quad \lim_{h \rightarrow 0} \frac{f(x_0+h) - f(x_0) - L_{x_0}(h)}{h} = \lim_{h \rightarrow 0} \frac{\Delta f_{x_0}(h) - L_{x_0}(h)}{h} = 0.$$

In tale ipotesi la funzione L_{x_0} si chiama differenziale della funzione f nel punto x_0 , e si indica con il simbolo

$$(19) \quad df_{x_0}.$$

Si noti che il differenziale di una funzione in un punto è una funzione (lineare), non un numero. Inoltre, posto $\omega(h) = \Delta f_{x_0}(h) - df_{x_0}(h)$, si ha

$$(20) \quad \Delta f_{x_0}(h) = df_{x_0}(h) + \omega(h),$$

e si deduce dalla definizione precedente che *il differenziale di una funzione in un punto x_0 , relativamente all'incremento h della variabile, approssima l'incremento della funzione nelle stesse condizioni a meno di un infinitesimo di ordine superiore ad h .*

Se teniamo di nuovo conto della (6), possiamo dire che se f è differenziabile in un punto x_0 , si ha

$$(21) \quad f(x) = f(x_0) + df_{x_0}(x - x_0) + \omega(x - x_0),$$

dove ω è una funzione infinitesima, al tendere di x a x_0 , di ordine superiore rispetto a $x - x_0$.

La situazione che abbiamo riscontrato per la funzione $f(x) = x^2$ non è eccezionale: in effetti, per le funzioni di una variabile reale, si dimostra il seguente teorema.

Teorema 3. Una funzione $f: I \rightarrow \mathbb{R}$ è differenziabile in un punto $x_0 \in I$ se e solo se è derivabile in x_0 e, in questo caso, si ha⁽²⁾

$$(22) \quad L_{x_0}(x - x_0) = df_{x_0}(x - x_0) = f'(x_0)(x - x_0).$$

Ne segue che la formula (21) si può riscrivere nella forma

$$(23) \quad f(x) = f(x_0) + f'(x_0)(x - x_0) + \omega(x - x_0).$$

È proprio in conseguenza della formula (23) che la retta r di equazione $y = f(x_0) + f'(x_0)(x - x_0)$ prende il nome di *retta tangente* al grafico della funzione f , in corrispondenza al punto x_0 : la formula infatti afferma che la funzione f può essere approssimata con la retta r a meno di un infinitesimo di ordine superiore rispetto a $x - x_0$, ovvero che, almeno in un intorno di x_0 , i grafici della funzione f e della retta r sono “quasi coincidenti”. Naturalmente bisogna sempre tenere conto del significato di infinitesimo o di infinitesimo di ordine superiore: l'approssimazione è buona in un “opportuno intorno” del punto x_0 , ma nulla si può affermare circa l'ampiezza di questo intorno. Può benissimo darsi che, su tutto un intervallo *prefissato* circondante il punto x_0 , ci siano altre approssimazioni lineari molto migliori “globalmente” di quella fornita “localmente” dalla retta tangente.

Si tenga ben presente che il differenziale di una funzione non ha come grafico la retta tangente: il differenziale è una funzione lineare dell'incremento h (e dunque ha come grafico una retta per l'origine); la retta tangente si ottiene, vedi la formula (21), calcolando il differenziale in $x - x_0$ e aggiungendo $f(x_0)$, ovvero traslando il grafico del differenziale in modo che il punto $(0, 0)$ vada a coincidere con il punto $(x_0, f(x_0))$.

È opportuno dare alla formula (23) una forma adatta ad essere generalizzata a funzioni di \mathbb{R}^n in \mathbb{R} (e poi anche a casi ancora più complessi). Per questo:

- indichiamo con \vec{x} il generico punto del dominio di f : per funzioni di una variabile il “vettore” \vec{x} avrà solo una componente, per due variabili ne avrà invece due: $\vec{x} = (x_1, x_2)$, ecc;
- detto \vec{i} il versore dell'asse x , chiamiamo *gradiente* della funzione f nel punto \vec{x}_0 il vettore (sempre a una componente!)

$$(24) \quad \nabla f(\vec{x}_0) = f'(\vec{x}_0)\vec{i}.$$

Se indichiamo il prodotto scalare (canonico) con “ \cdot ”, il differenziale della funzione si potrà scrivere come

$$(25) \quad df_{\vec{x}_0}(\vec{x} - \vec{x}_0) = \nabla f(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0)$$

e la formula (23) si riscriverà nella forma seguente:

$$(26) \quad f(\vec{x}) = f(\vec{x}_0) + \nabla f(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0) + \omega(\vec{x} - \vec{x}_0),$$

forma che ritroveremo intatta trattando il caso di più dimensioni.

Osserviamo che, con queste notazioni, l'equazione della retta tangente al grafico della funzione f nel punto $(x_0, f(x_0))$ si può scrivere in una delle forme

$$(27) \quad y = f(x_0) + df_{x_0}(x - x_0), \quad y = f(\vec{x}_0) + \nabla f_{x_0} \cdot (\vec{x} - \vec{x}_0).$$

La figura 4 seguente, dove ci siamo limitati al caso $h > 0$ per semplicità, illustra il significato geometrico del differenziale.

²Si noti, nella formula (22), una certa difficoltà di lettura dovuta alle notazioni: mentre in $df_{x_0}(x - x_0)$ le parentesi tonde racchiudono l'argomento della funzione df_{x_0} , in $f'(x_0)(x - x_0)$ la seconda coppia di parentesi serve a racchiudere il termine $(x - x_0)$ nel prodotto di due fattori.

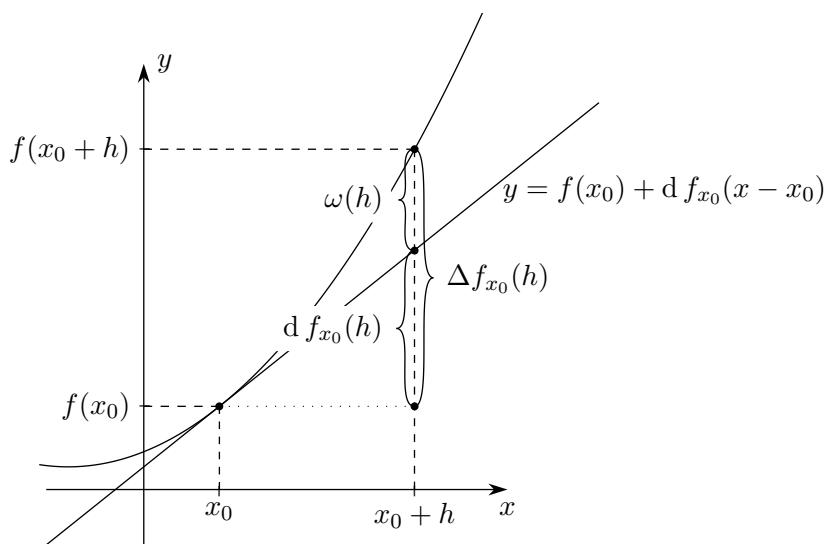


Figura 4 Differenziale e significato geometrico

3 Funzioni di \mathbb{R}^2 in \mathbb{R}

Sia ora $f: D \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ una funzione definita in un sottoinsieme (che supporremo aperto per semplicità) di \mathbb{R}^2 e a valori in \mathbb{R} . Sia $\vec{x}_0 = (x_{1,0}, x_{2,0})$ un punto di D e $f(\vec{x}_0)$ il corrispondente valore della funzione. Come nel caso di una variabile siamo interessati a valutare la variazione della funzione al passaggio da \vec{x}_0 a un punto generico \vec{x} situato nei pressi di \vec{x}_0 ; se poniamo

$$(28) \quad \vec{h} = \vec{x} - \vec{x}_0 = (h_1, h_2),$$

questa variazione sarà data da

$$(29) \quad \Delta f_{\vec{x}_0}(\vec{h}) = f(\vec{x}_0 + \vec{h}) - f(\vec{x}_0),$$

che è una funzione di \vec{h} , una volta fissato \vec{x}_0 . Anche in questo caso vogliamo verificare se è possibile approssimare l'incremento con una funzione lineare, ovvero una funzione avente come grafico un piano passante per l'origine. In perfetta analogia con la definizione 2, possiamo dare la definizione seguente di funzione differenziabile.

Definizione 4. Sia $f: D \rightarrow \mathbb{R}$ una funzione definita in un aperto D di \mathbb{R}^2 . La funzione si dice differenziabile nel punto $\vec{x}_0 \in D$ se esiste una funzione lineare $L_{\vec{x}_0}: \mathbb{R}^2 \rightarrow \mathbb{R}$ tale che

$$(30) \quad \lim_{(h_1, h_2) \rightarrow (0,0)} \frac{f(\vec{x}_0 + \vec{h}) - f(\vec{x}_0) - L_{\vec{x}_0}(\vec{h})}{\|\vec{h}\|} = \lim_{(h_1, h_2) \rightarrow (0,0)} \frac{\Delta f_{\vec{x}_0}(\vec{h}) - L_{\vec{x}_0}(\vec{h})}{\|\vec{h}\|} = 0$$

In tale ipotesi la funzione $L_{\vec{x}_0}$ si chiama differenziale della funzione f nel punto \vec{x}_0 , e si indica con il simbolo

$$(31) \quad df_{\vec{x}_0}.$$

Anche in questo caso si può notare che il differenziale di una funzione in un punto è una funzione (lineare), non un numero, e precisamente, trattandosi di una funzione lineare di \mathbb{R}^2 in \mathbb{R} , una funzione del tipo

$$(32) \quad L_{\vec{x}_0}(\vec{h}) = L_{\vec{x}_0}(h_1, h_2) = df_{\vec{x}_0}(h_1, h_2) = a_1 h_1 + a_2 h_2,$$

dove i due coefficienti a_1, a_2 dipendono da $\vec{x}_0 = (x_{1,0}, x_{2,0})$. Naturalmente anche in questo caso la definizione data implicherà che

$$(33) \quad \Delta f_{\vec{x}_0}(\vec{h}) = df_{\vec{x}_0}(\vec{h}) + \omega(\vec{h}),$$

ovvero, tenendo conto che $\vec{h} = \vec{x} - \vec{x}_0$,

$$(34) \quad f(\vec{x}) = f(\vec{x}_0) + df_{\vec{x}_0}(\vec{x} - \vec{x}_0) + \omega(\vec{x} - \vec{x}_0).$$

La funzione ω è infinitesima, per $\vec{x} \rightarrow \vec{x}_0$, di ordine superiore rispetto a $\|\vec{x} - \vec{x}_0\|$, cioè rispetto alla distanza tra \vec{x} e \vec{x}_0 .

Questa volta la situazione analitica è più complessa che non nel caso delle funzioni di una variabile e può tranquillamente succedere che la funzione f abbia, in \vec{x}_0 , entrambe le derivate parziali, o addirittura tutte le derivate direzionali, senza essere differenziabile. Del resto, come è ben noto, l'esistenza delle derivate parziali e, purtroppo, anche di tutte le derivate direzionali in un punto non è nemmeno condizione sufficiente per la continuità della funzione. Vale però il seguente teorema, che esprime in sostanza una condizione necessaria per la differenziabilità.

Teorema 5. *Se una funzione $f: D \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ è differenziabile in un punto \vec{x}_0 , allora è continua e parzialmente derivabile in \vec{x}_0 e inoltre si ha*

$$(35) \quad df_{\vec{x}_0}(\vec{x} - \vec{x}_0) = \frac{\partial f}{\partial x_1}(\vec{x}_0)(x_1 - x_{1,0}) + \frac{\partial f}{\partial x_2}(\vec{x}_0)(x_2 - x_{2,0}).$$

Ridefinendo il gradiente della funzione f come estensione della definizione data nel caso di una variabile, si ottiene:

$$(36) \quad \nabla f(\vec{x}_0) = \frac{\partial f}{\partial x_1}(\vec{x}_0)\vec{i} + \frac{\partial f}{\partial x_2}(\vec{x}_0)\vec{j}.$$

Dunque il differenziale si potrà scrivere come

$$(37) \quad df_{\vec{x}_0}(\vec{x} - \vec{x}_0) = \nabla f(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0)$$

e la formula (34) si riscriverà nella forma

$$(38) \quad f(\vec{x}) = f(\vec{x}_0) + \frac{\partial f}{\partial x_1}(\vec{x}_0)(x_1 - x_{1,0}) + \frac{\partial f}{\partial x_2}(\vec{x}_0)(x_2 - x_{2,0}) + \omega(\vec{x} - \vec{x}_0),$$

oppure

$$(39) \quad f(\vec{x}) = f(\vec{x}_0) + \nabla f(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0) + \omega(\vec{x} - \vec{x}_0).$$

Per gli stessi motivi già segnalati nel caso delle funzioni di una variabile, il piano di equazione

$$(40) \quad y = f(\vec{x}_0) + \frac{\partial f}{\partial x_1}(\vec{x}_0)(x_1 - x_{1,0}) + \frac{\partial f}{\partial x_2}(\vec{x}_0)(x_2 - x_{2,0})$$

è il *piano tangente* al grafico della funzione f , nel punto x_0 .

È importante il fatto che se una funzione è differenziabile in un punto \vec{x}_0 allora è ivi derivabile secondo una qualunque direzione orientata e, detto $\vec{u} = (l, m)$ il versore della direzione, si ha

$$(41) \quad \frac{\partial f}{\partial \vec{u}}(\vec{x}_0) = \nabla f(\vec{x}_0) \cdot \vec{u} = \frac{\partial f}{\partial x_1}(\vec{x}_0)l + \frac{\partial f}{\partial x_2}(\vec{x}_0)m.$$

A scanso di equivoci è opportuno ricordare (ancora una volta!) che non vale il viceversa: una funzione può essere derivabile lungo ogni direzione senza essere differenziabile, e in questo caso la derivata direzionale *non* si potrà esprimere con la formula (41).

La figura 5 seguente illustra il significato geometrico del differenziale per una funzione di due variabili. Nella figura sono rappresentati:

- una parte del dominio D della funzione nel piano Ox_1x_2 , precisamente il rettangolo R di vertici opposti \vec{x}_0 e $\vec{x} = \vec{x}_0 + \vec{h}$;
- la parte di piano $y = f(\vec{x}_0)$ sovrastante il rettangolo R ;
- la parte di piano tangente sovrastante il rettangolo R : $y = f(\vec{x}_0) + \nabla f(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0)$;
- il lembo della superficie-grafico della funzione f , sovrastante il rettangolo R .

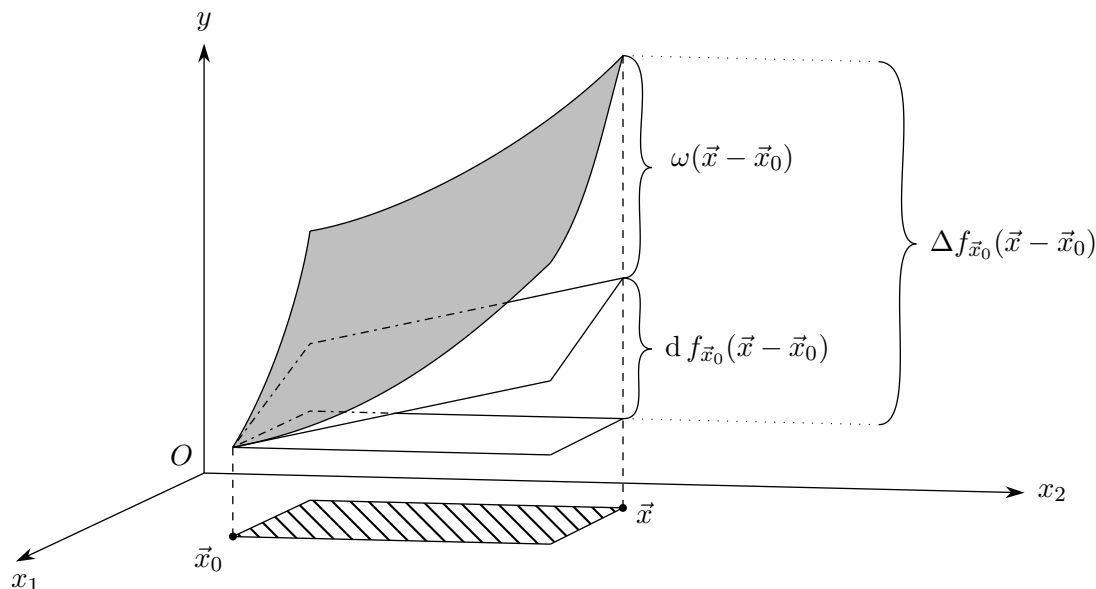


Figura 5 Differenziale in due variabili e significato geometrico

Si può verificare la sostanziale identità tra le immagini riportate nelle figure 4 e 5, e in particolare il fatto che il piano tangente costituisce una approssimazione del grafico della funzione, nei pressi del punto \vec{x}_0 .

4 Funzioni di \mathbb{R}^n in \mathbb{R}

Mentre il passaggio dalle funzioni di una variabile a quelle di due variabili è abbastanza delicato, in quanto l'esistenza delle derivate parziali non garantisce, da sola, la differenziabilità di una funzione, il passaggio a un numero di variabili maggiore di 2 ormai è solo questione formale: nulla cambia nella sostanza. Si tratterà solo di tenere conto del fatto che \vec{x} , \vec{x}_0 , \vec{h} avranno n componenti, anziché solo due. Naturalmente non si avrà più a disposizione una rappresentazione cartesiana significativa come quella proposta nella figura 5, e il concetto di *piano tangente* sarà sostituito da quello di *iperpiano tangente*, sempre senza alcuna possibilità di interpretazione geometrica elementare.

Rimane il fatto essenziale (ed è questo che serve nelle applicazioni, al di là della visualizzazione grafica!), che l'incremento di una funzione si può approssimare, in un intorno di un punto dove la funzione è differenziabile, con una funzione lineare, il differenziale $df_{\vec{x}_0}$. Approssimazioni di ordine più elevato (polinomiali di grado due, tre, ecc.) sono oggetto della formula di Taylor.

Riscriviamo, a titolo d'esempio, le formule che si riferiscono al caso di una funzione di \mathbb{R}^3 in \mathbb{R} , in un punto \vec{x}_0 .

- Gradiente di una funzione: $\nabla f(\vec{x}_0) = \frac{\partial f}{\partial x_1}(\vec{x}_0)\vec{i} + \frac{\partial f}{\partial x_2}(\vec{x}_0)\vec{j} + \frac{\partial f}{\partial x_3}(\vec{x}_0)\vec{k}$.

— Differenziale di una funzione:

$$df_{\vec{x}_0}(\vec{x}-\vec{x}_0) = \nabla f(\vec{x}_0) \cdot (\vec{x}-\vec{x}_0) = \frac{\partial f}{\partial x_1}(\vec{x}_0)(x_1-x_{1,0}) + \frac{\partial f}{\partial x_2}(\vec{x}_0)(x_2-x_{2,0}) + \frac{\partial f}{\partial x_3}(\vec{x}_0)(x_3-x_{3,0}).$$

— Iperpiano tangente: $y = f(\vec{x}_0) + \nabla f(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0)$.

— Derivata secondo la direzione del versore $\vec{u} = (l, m, n)$:

$$\frac{\partial f}{\partial \vec{u}}(\vec{x}_0) = \nabla f(\vec{x}_0) \cdot \vec{u} = \frac{\partial f}{\partial x_1}(\vec{x}_0)l + \frac{\partial f}{\partial x_2}(\vec{x}_0)m + \frac{\partial f}{\partial x_3}(\vec{x}_0)n.$$

Ricordiamo, per completezza, una condizione *sufficiente* per la differenziabilità di una funzione di \mathbb{R}^n in \mathbb{R} , ($n \geq 2$).

Teorema 6 (del differenziale totale). *Sia $f: D \subset \mathbb{R}^n \rightarrow \mathbb{R}$, una funzione definita in un (aperto) D di \mathbb{R}^n . Se f è parzialmente derivabile in un intorno I di un punto \vec{x}_0 e le derivate parziali sono continue in \vec{x}_0 , allora la funzione è differenziabile in \vec{x}_0 .*

Si noti come questo teorema richieda la continuità delle funzioni derivate parziali in \vec{x}_0 , e non la semplice esistenza delle derivate parziali: mentre l'esistenza della derivate parziali concerne solo la restrizione della funzione a una direzione, passante per \vec{x}_0 , parallela agli assi, la continuità delle derivate stesse concerne tutto un intorno n -dimensionale del punto \vec{x}_0 , e questo è in accordo con il fatto che la differenziabilità è proprio una questione che coinvolge tutto un intorno n -dimensionale del punto \vec{x}_0 .

5 La scrittura tradizionale

Tradizionalmente le formule relative al differenziale vengono scritte in una forma molto significativa per le applicazioni, ma che va presa con molta cautela. Ne facciamo solo un breve cenno.

Il caso delle funzioni di una variabile

Cominciamo con il caso delle funzioni di una variabile e consideriamo la funzione

$$(42) \quad g: \mathbb{R} \rightarrow \mathbb{R}, \quad g(x) = x,$$

cioè la funzione identica di \mathbb{R} in \mathbb{R} .

Poiché $g'(x_0) = 1, \forall x_0 \in \mathbb{R}$, il differenziale della funzione g sarà lo stesso (cioè la stessa funzione) in ogni punto, e precisamente la funzione $L(h) = h$, ovvero

$$(43) \quad dg_{x_0}(h) = h \quad \Rightarrow \quad dg_{x_0}(x - x_0) = x - x_0.$$

È abitudine indicare la funzione g semplicemente con x ⁽³⁾. Tenendo anche conto che il differenziale non dipende da x_0 , si omette di esplicitare questa dipendenza, e anziché $dg_{x_0}(x - x_0)$ si scrive semplicemente $dx(x - x_0)$ (“ dx ” calcolato in $x - x_0$). Anzi, siccome il differenziale si calcola naturalmente *sempre* in $x - x_0$, cioè nell’incremento, si evita anche di esplicitare il fatto che il dx è calcolato in $x - x_0$ e la (43) si scrive semplicemente

$$(44) \quad dx = x - x_0.$$

³Si tratta di una scrittura che può dar luogo ad equivoci, ma che può essere precisata con rigore. Come sempre nella lettura di una formula matematica occorre avere presente le convenzioni adottate (e tenere presente che non si può sempre scrivere tutto con la più assoluta precisione, altrimenti certe formule diventerebbero chilometriche!).

A questo punto anche nel differenziale di una generica funzione f di una variabile si omette di esplicitare il fatto che il differenziale va calcolato in $x - x_0$, e si scrive semplicemente

$$(45) \quad df_{x_0} = f'(x_0) dx.$$

È (anche) la formula (45) che dà origine al tradizionale simbolo per la derivata di una funzione:

$$(46) \quad f'(x_0) = \frac{df}{dx}(x_0).$$

Il caso delle funzioni di più variabili

Considerazioni simili si possono fare per funzioni reali di più variabili reali. Occorrerà sempre tenere ben presente che la differenziabilità è, ora, una condizione più restrittiva che non la semplice esistenza delle derivate parziali.

Trattiamo esplicitamente il caso di due variabili, il caso generale è assolutamente identico. Consideriamo le due funzioni

$$(47) \quad \pi_1(\vec{x}) = x_1, \quad \pi_2(\vec{x}) = x_2,$$

dove i nomi π_1 e π_2 sono stati scelti perché le due funzioni sostanzialmente “proiettano” il generico punto \vec{x} del piano Ox_1x_2 sull’asse x_1 e x_2 rispettivamente, fornendo la prima e la seconda componente di \vec{x} . È immediato constatare che, per ogni \vec{x}_0 ,

$$(48) \quad \frac{\partial \pi_1}{\partial x_1}(\vec{x}_0) = 1, \quad \frac{\partial \pi_1}{\partial x_2}(\vec{x}_0) = 0, \quad \text{mentre} \quad \frac{\partial \pi_2}{\partial x_1}(\vec{x}_0) = 0, \quad \frac{\partial \pi_2}{\partial x_2}(\vec{x}_0) = 1.$$

Dunque

$$(49) \quad d\pi_{1,\vec{x}_0}(\vec{x} - \vec{x}_0) = x_1 - x_{1,0}, \quad d\pi_{2,\vec{x}_0}(\vec{x} - \vec{x}_0) = x_2 - x_{2,0},$$

da cui, con le stesse convenzioni utilizzate nel caso di una variabile,

$$(50) \quad dx_1 = x_1 - x_{1,0}, \quad dx_2 = x_2 - x_{2,0},$$

e, infine,

$$(51) \quad df_{\vec{x}_0} = \frac{\partial f}{\partial x_1}(\vec{x}_0) dx_1 + \frac{\partial f}{\partial x_2}(\vec{x}_0) dx_2.$$

6 Funzioni di \mathbb{R}^n in \mathbb{R}^m

Le funzioni di \mathbb{R}^n in \mathbb{R}^m non presentano sostanziali novità rispetto ai casi precedenti, in quanto si possono pensare come m -uple di funzioni di \mathbb{R}^n in \mathbb{R} , dette funzioni componenti: tutte le considerazioni fatte si applicheranno alle singole funzioni componenti.

La cosa significativa da segnalare è che, avendo m funzioni di \mathbb{R}^n in \mathbb{R} , avremo m gradienti, ciascuno con n componenti: è ovvio che essi potranno essere scritti in una matrice, J , a m righe e n colonne, come segue.

$$(52) \quad J(\vec{x}_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix},$$

dove tutte le derivate vanno calcolate in \vec{x}_0 . La matrice così costruita si chiama *matrice jacobiana* e le formule che legano il differenziale al gradiente saranno sostituite dalla formula

$$(53) \quad df_{\vec{x}_0}(\vec{x} - \vec{x}_0) = J(\vec{x}_0)(\vec{x} - \vec{x}_0).$$

Continueranno a valere tutte le considerazioni che abbiamo già fatto in precedenza, compreso il *Teorema del differenziale totale* che, per maggiore chiarezza, riformuliamo qui di seguito.

Teorema 7 (del differenziale totale). *Sia $f: D \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, una funzione definita in un (aperto) D di \mathbb{R}^n . Se tutte le componenti (f_1, f_2, \dots, f_m) di f sono parzialmente derivabili in un intorno I di un punto \vec{x}_0 e le derivate parziali sono continue in \vec{x}_0 , allora la funzione è differenziabile in \vec{x}_0 , cioè esiste una funzione lineare $L_{\vec{x}_0}: \mathbb{R}^n \rightarrow \mathbb{R}^m$, tale che*

$$(54) \quad \lim_{\vec{h} \rightarrow \vec{0}} \frac{f(\vec{x}_0 + \vec{h}) - f(\vec{x}_0) - L_{\vec{x}_0}(\vec{h})}{\|\vec{h}\|} = \lim_{\vec{h} \rightarrow \vec{0}} \frac{\Delta f_{\vec{x}_0}(\vec{h}) - L_{\vec{x}_0}(\vec{h})}{\|\vec{h}\|} = 0.$$

È opportuno segnalare che nel caso di funzioni $\mathbb{R} \rightarrow \mathbb{R}^m$, poiché le singole componenti sono funzioni di una sola variabile, varranno tutte le proprietà specifiche di questo caso, in particolare la differenziabilità coinciderà con la derivabilità (ordinaria perché si tratta di funzioni di una sola variabile) delle m componenti della funzione. Si tratta di un caso particolarmente importante: le curve di \mathbb{R}^m (in particolare di \mathbb{R}^2 e \mathbb{R}^3) sono proprio funzioni di questo tipo: dunque la differenziabilità è essenzialmente diversa dalla derivabilità quando il dominio è multidimensionale.

7 Algebra dei differenziali

Concludiamo questa serie di osservazioni sul concetto di differenziale ricordando i seguenti risultati.

1. Siano f e g due funzioni di \mathbb{R}^n in \mathbb{R}^m , entrambi differenziabili in \vec{x}_0 . Allora $f+g$ è differenziabile in \vec{x}_0 e si ha

$$(55) \quad d(f+g)_{\vec{x}_0}(\vec{x} - \vec{x}_0) = df_{\vec{x}_0}(\vec{x} - \vec{x}_0) + dg_{\vec{x}_0}(\vec{x} - \vec{x}_0).$$

2. Siano $f: A \subset \mathbb{R}^n \rightarrow \mathbb{R}^p$ una funzione differenziabile in $\vec{x}_0 \in A$ e $g: B \subset \mathbb{R}^p \rightarrow \mathbb{R}^m$, una funzione differenziabile in $\vec{y}_0 = f(\vec{x}_0) \in B$; sia inoltre $h = g \circ f: A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ la funzione composta di f e g . Allora h è differenziabile in \vec{x}_0 e si ha

$$(56) \quad dh_{\vec{x}_0}(\vec{x} - \vec{x}_0) = dg_{\vec{y}_0}(\vec{y} - \vec{y}_0) \circ df_{\vec{x}_0}(\vec{x} - \vec{x}_0).$$

In termini di matrici jacobiane la formula (56) si scrive

$$(57) \quad J_h(\vec{x}_0) = J_g(\vec{y}_0)J_f(\vec{x}_0),$$

con ovvio significato dei simboli.

È in particolare la seconda delle due proprietà elencate che giustifica l'importanza del concetto di differenziale, come estensione naturale a più dimensioni del concetto di derivata: mentre la composta di due funzioni (di più variabili!) derivabili (magari lungo ogni direzione), può non essere affatto derivabile (anzi addirittura nemmeno continua), la composta di due funzioni differenziabili è differenziabile, esattamente come succedeva per la derivabilità nel caso di funzioni di una variabile.