ELEMENTI DI CALCOLO DIFFERENZIALE IN PIÙ VARIABILI. PARTE I

FAUSTO FERRARI

Materiale propedeutico alle lezioni di Analisi Matematica per i corsi di Laurea in Ingegneria Energetica e Meccanica N-Z dell'Università di Bologna. Anno Accademico 2003/2004

1. Cenni di topologia

Per poter estendere i risultati ottenuti per funzioni di una variabile al caso di funzioni di più variabili occorre innanzitutto avere uno strumento matematico di valutazione delle distanze tra i punti di $\mathbb{R}^n$. D'altra parte, per gli scopi che il nostro corso si prefigge, avendo a che fare con una struttura molto ricca di proprietà come $\mathbb{R}^n$ possiamo definire una funzione distanza a partire dal prodotto scalare definito su $\mathbb{R}^n$.

Definizione 1.1. *(Prodotto scalare) Sia $\{e_1, \ldots, e_n\}$ la base canonica di $\mathbb{R}^n$ Per ogni $x, y \in \mathbb{R}^n$, $\sum_{i=1}^n x_i e_i$ e $\sum_{i=1}^n y_i e_i$. Definiamo il prodotto scale su $\mathbb{R}^n$ la seguente funzione $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$*

$$\langle x, y \rangle = \sum_{i=1}^n x_i y_i.$$

Propoposizione 1.1. *Il prodotto scalare è:*

i) *bilineare, cioè per ogni $y \in \mathbb{R}^n$ la funzione $\langle \cdot, y \rangle : \mathbb{R}^n \to \mathbb{R}$*

$$x \to \langle x, y \rangle$$

è lineare. Analogamente per ogni $x \in \mathbb{R}^n$ la funzione $\langle x, \cdot \rangle : \mathbb{R}^n \to \mathbb{R}$

$$y \to \langle x, y \rangle$$

è lineare;

ii) *simmetrico, cioè per ogni $x, y \in \mathbb{R}^n$*

$$\langle x, y \rangle = \langle y, x \rangle$$

A partire dal prodotto scalare definiamo la funzione norma associata al prodotto scalare nel modo seguente.

Definizione 1.2. *Definiamo norma associata al prodotto scalare $\langle \cdot, \cdot \rangle$ la seguente funzione $||\cdot|| : \mathbb{R}^n \to \mathbb{R}^+ \cup \{0\}$ così definita*

$$||x|| = \sqrt{\langle x, x \rangle}.$$

La norma di un vettore è "l'equivalente" del valore assoluto di un numero ed è tramite la norma che possiamo considerare una famiglia di insiemi di $\mathbb{R}^n$ di inportanza fondamentale: le palle di $\mathbb{R}^n$.

Definizione 1.3. *Per ogni punto $x_0 \in RR^n$ e per ogni numero positivo fissato chiamiamo palla ci centro x_0 e raggio r il seguente insieme:*

$$B(x_0, r) = \{x \in \mathbb{R}^n : \ ||x - x_0|| < r\}.$$

Date: 14/02/2004.

Nel caso di dimensione 2 l'equazione $||x - x_0|| = r$ individua una circonferenza di raggio r e centro $x_0 = (x_{01}, x_{02}) \in \mathbb{R}^2$. Quindi in questo caso si può parlare di disco anziché di palla. Nel caso di $\mathbb{R}^3$, $||x - x_0|| = r$ individua una sfera di raggio r e centro $x_0 = (x_{01}, x_{02}, x_{03}) \in \mathbb{R}^3$. Una delle proprietà fondamentali della norma è che soddisfa la *disuguaglianza triangolare* ed è omogenea di grado 1

Propoposizione 1.2. *Per ogni $x, y \in \mathbb{R}^n$*

a) *Per ogni $x, y \in \mathbb{R}^n$,*
$$||x + y|| \leq ||x|| + ||y|| \quad \textit{(disuguaglianza triangolare)}$$
b) *Per ogni $c \in \mathbb{R}$ e per ogni $x \in \mathbb{R}^n$,*
$$||cx|| = | c | \, ||x||$$
c) *Per ogni $x \in \mathbb{R}^n$,*
$$||x|| = || - x||.$$

La norma gode quindi sostanzialmente delle stesse proprietà del valore assoluto. Per concludere questo sommario paragone tra norma e valore assoluto notiamo la seguente importante differenza tra le due funzioni

Teorema 1.1. *(Disuguaglianza di Cauchy Schwarz) Per ogni $x, y \in \mathbb{R}^n$*
$$| \langle x, y \rangle | \leq ||x|| \cdot ||y||.$$

Introduciamo alcuni elementi di topologia.

Definizione 1.4. *Sia $\Omega \subset \mathbb{R}^n$ un insieme e $x_0 \in \mathbb{R}^n$.*

i) *x_0 è punto interno ad Ω se esiste $r > 0$ tale che $B(x_0, r) \subset \Omega$*
ii) *x_0 è punto esterno ad Ω se esiste $r > 0$ tale che $B(x_0, r) \subset \Omega'$, dove Ω' è l'insieme complementare di Ω in $\mathbb{R}^n$, cioè $\Omega' = \mathbb{R}^n \setminus \Omega$.*
iii) *x_0 è punto di frontiera di Ω se per ogni $r > 0$*
$$B(x_0, r) \cap \Omega \neq \emptyset \text{ e } B(x_0, r) \cap \Omega' \neq \emptyset$$
.

Diamo la seguente definizione di interno di un insieme

Definizione 1.5. *L'interno di un insieme $\Omega \subset \mathbb{R}^n$ è l'insieme di tutti i punti inrterni di Ω. Denoteremo l'insieme interno di Ω con il simbolo int(Ω).*

Definizione 1.6. *(Def. di aperto) Sia $\Omega \subset \mathbb{R}^n$ un insieme. Diremo che Ω è aperto se int$(\Omega) = \Omega$.*

Definizione 1.7. *(Def. di chiuso) Sia $\Omega \subset \mathbb{R}^n$ un insieme. Diremo che Ω è chiuso se $\mathbb{R}^n \setminus \Omega$ è aperto.*

Definizione 1.8. *(Def. di frontiera) Sia $\Omega \subset \mathbb{R}^n$ un insieme. La frontiera di Ω è l'insieme di tutti i punti di $\mathbb{R}^n$ che sono punti di frontiera di Ω. Denoteremo questo insieme con il simbolo Fr(Ω).*

Definizione 1.9. *(Def. di chiusura) Sia $\Omega \subset \mathbb{R}^n$ un insieme. La chiusura di Ω è il seguente insieme*
$$\bar{\Omega} = \Omega \cup Fr(\Omega).$$

Si può dimostrare che $\Omega \subset \mathbb{R}^n$ è chiuso se e solo se $\bar{\Omega} = \Omega$.

La varietà degli insiemi è tale per cui occorre classificare i punti degli insiemi per i quali ha senso introdurre le nozioni di limite e derivata. Analogamente a quanto abbiamo già visto nel corso di Analasi matematica L-A le successioni giocano un ruolo fondamentale.

Definizione 1.10. *Sia $\Omega \subset \mathbb{R}^n$ e $x_0 \in \mathbb{R}^n$. Diremo che x_0 è un punto di accumulazione per Ω se per ogni $r > 0$*

$$B(x_0, r) \cap (\Omega \setminus \{x_0\}) \neq \emptyset.$$

L'insieme di tutti i punti di accumulazione di Ω è il derivato di Ω e sarà denotato con $D(\Omega)$.

Definizione 1.11. *(Definizione di successione convergente) Diremo che una successione $(x_j)_{j\in N} \subset \mathbb{R}^n$ a $x_0 \in \mathbb{R}^n$ se:*

$$\textit{per ogni } \epsilon > 0 \textit{ esiste } \bar{n} \in N : ||x_j - x_0|| < \epsilon, \textit{per ogni } n \in N, n > \bar{n}.$$

In tal caso diremo che la successione è convergente a x_0 e scriveremo $x_j \to x_0$, $j \to \infty$.

Attenzione alle notazioni! I punti $x_j \in \mathbb{R}^n$, quindi le coordinate del punto sono n e

$$x_j = (x_j^{(1)}, \dots, x_j^{(n)}).$$

Teorema 1.2. *Una successione $(x_j)_{j\in N} \subset \mathbb{R}^n$ converge a $x_0 \in \mathbb{R}^n$ se e solo se per ogni $k \in \{1, \dots, n\}$, la successione $(x_j^{(k)})_{j\in N} \subset \mathbb{R}$ converge in $\mathbb{R}$ a x_0^k.*

La prova di questo risultato è la seguente. Se $x_j \to x_0$, $j \to \infty$, allora

$$\text{per ogni } \epsilon > 0 \text{ esiste } \bar{n} \in N : ||x_j - x_0|| < \epsilon, \text{per ogni } n \in N, n > \bar{n}.$$

In particolare per ogni $k \in \{1, \dots, n\}$, per semplicità prendiamo $k = 1$, $(x_j^{(1)})_{j\in N} \subset \mathbb{R}$ è una successione convergente a $x_0^{(1)}$, perché per ogni $\epsilon > 0$ esiste $\bar{n} \in N$, esattamente lo stesso indice legato allipotesi di esistenza del limite della successione in $\mathbb{R}^n$, tale che

$$|x_j^{(1)} - x_0^{(1)}| \leq \sqrt{\sum_{p=1}^{n} (x_j^p - x_0^p)^2} = ||x_j - x_0|| < \epsilon$$

per ogni $n \in N, n > \bar{n}$. Cioè la successione $x_j^{(1)} \to x_0^1$, per $j \to \infty$, inoltre $x_j^{(k)} \to x_0^k$, per $j \to \infty$ per $k \in \{1, \dots n\}$. Vediamo l'implicazione opposta. Supponiamo cioè che per ogni $k \in \{1, \dots n\}$, $x_j^{(k)} \to x_0^k$, per $j \to \infty$. Dalla definizione di limite di successione in $\mathbb{R}$ segue che per ogni $k \in \{1, \dots n\}$ e per ogni $\epsilon > 0$ esistono $\bar{n}_k \in N$:

$$|x_j^{(k)} - x_0^{(k)}| < \epsilon,$$

per ogni $j \in N$, $j > \bar{n}_k$. Fissiamo $q = \max_{1\leq k\leq n, k\in N} \bar{n}_k$. Avremo che

$$\sqrt{\sum_{p=1}^{n} (x_j^p - x_0^p)^2} = ||x_j - x_0|| < n\epsilon,$$

per ogni $j \in N$, $j > q$. Questo significa che $x_j \to x_0$. E' inoltre immediata la dimostrazione del seguente Teorema.

Teorema 1.3. *Per ogni insieme $\Omega \subset \mathbb{R}^n$, $x_0 \in D(\Omega)$ se e solo se esiste una successione di punti $(x_j)_{j\in N} \subset (\Omega \setminus \{x_0\})$ tale che*

$$x_n \to x_0.$$

Per ragioni di semplicità espositiva formuleremo i seguenti risultati in $\mathbb{R}^2$, quindi le componenti sono solo 2 cioè $n = 2$

Definizione 1.12. *Sia $\Omega \subset \mathbb{R}^2$, $x_0 \in D(\Omega)$ e $f : \Omega \to \mathbb{R}$ una funzione. Diremo che la funzione f tende a $\alpha \in \mathbb{R}$ per $x \to x_0$ se:*

per ogni successione $(x_j)_{j\in N} \subset (\Omega \setminus \{x_0\})$, tale che $x_j \to x_0$, $j \to \infty$, allora $\lim_{j\to\infty} f(x_j) = \alpha$.
In tal caso diremo che esiste il limite della funzione f per $x \to x_0$ e scriveremo $f(x) \to \alpha$, $j \to \infty$.

2. Funzioni continue

Definizione 2.1. *Sia $\Omega \subset \mathbb{R}^2$, $x_0 \in \Omega$ e $f : \Omega \to \mathbb{R}$ una funzione. Diremo che la funzione f è continua in x_0 se:*
per ogni successione $(x_j)_{j\in N} \subset \Omega$, tale che $x_j \to x_0$, $j \to \infty$, allora $\lim_{j\to\infty} f(x_j) = f(x_0)$.

Parleremo di funzione continua, in assenza di ulteriori precisazioni, sottointendendo che è continua in ogni punto del suo dominio.

Definizione 2.2. *(derivata secondo un vettore) Data $f : \Omega \to \mathbb{R}$ funzione, $\Omega \subset \mathbb{R}^2$ e $x_0 \in int(\Omega)$. Sia poi $a \in \mathbb{R}^2$ tale che $||a|| = 1$. Diremo che f è derivabile in x_0 rispetto al vettore a se esiste in $\mathbb{R}$ il seguente limite $\alpha \in \mathbb{R}$*

$$\lim_{t\to 0^+} \frac{f(x_0 + ta) - f(x_0)}{t} = \alpha.$$

Denoteremo tale numero α con il simbolo $\frac{\partial f(x_0)}{\partial a}$.

La precedente definizione non copre in modo soddisfacente il ruolo che vorremmo che avesse la derivata per funzioni in più variabili. Infatti il seguente esempio prova che esistono funzioni derivabili lungo ogni vettore in un punto senza che queste siano continue.

ESEMPIO $f : \mathbb{R}^2 \to \mathbb{R}$

$$f(x,y) = \begin{cases} \frac{2x^2 y}{x^4+y^2}, & \text{se } (x,y) \neq (0,0) \\ 0, & \text{se } (x,y) = (0,0). \end{cases}$$

ammette derivata rispetto ad ogni vettore in $(0,0)$, tuttavia non è continua in $(0,0)$, perché la successione $(\frac{1}{n}, \frac{1}{n^2})_{n\in N}$ valutata per mezzo di f è la successione

$$f(\frac{1}{n}, \frac{1}{n^2}) = 1$$

che ovviamente non converge a $f(0,0) = 0$.

Le precedenti definizioni di limite, di funzione continua e di derivata rispetto ad un vettore si estendono in modo naturale al caso in cui la funzione sia definita su di un sottoinsieme di $\mathbb{R}^n$ con $n > 2$.

Definizione 2.3. *Diremo che $\Omega \subset \mathbb{R}^2$ è limitato se esiste una costante positiva L tale che per ogni $x \in \Omega$*

$$||x|| < L.$$

Per esempio una successione $(x_j)_{j\in N} \subset \mathbb{R}^2$ è limitata se l'insieme $\{x_j : \jmath \in N\}$ è limitato.

Teorema 2.1. *Ogni successione limitata di $\mathbb{R}^2$ ammette una sottosuccessione convergente.*

Lemma 2.1. *Sia B un sottoinsieme chiuso di $\mathbb{R}^2$, e $(x_j)_{j\in N} \subset B$ una successione convergente a x. Allora $x \in B$.*

Definizione 2.4. *Chiameremo insieme compatto ogni insieme chiuso e limitato.*

Lemma 2.2. *Sia $K \subset \mathbb{R}^2$. Da ogni successione in K si estrae una sottosuccessione convergente in K.*

Introduciamo alcuni simboli nuovi. Sia $f : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$ una funzione. Scriveremo

$$\sup_\Omega f = \sup\{f(\Omega)\}; \ \inf_\Omega f = \inf\{f(\Omega)\}$$

$$\max_\Omega f = \max\{f(\Omega)\}; \ \min_\Omega f = \min\{f(\Omega)\}$$

Si dice che la funzione f è dotata di massimo (minimo) se l'insieme $f(\Omega)$ è dotato di massimo (minimo).

Teorema 2.2. *(Weierstrass) Sia $f : K \subset \mathbb{R}^2 \to \mathbb{R}$ una funzione continua in K e K un compatto di $\mathbb{R}^2$, allora f ̀dotata di massimo e minimo.*

La nozione di limite continuità di una funzione vettoriale in un punto è data per componenti.

Definizione 2.5. *Sia $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^p$ una funzione e $x_0 \in \Omega$.*

1) *Sia $x_0 \in D(\Omega)$; diremo che esiste il limite di f per $x \to x_0$ se per ogni $j \in \{1, \cdots, p\}$ esiste il limite di f_j $x \to x_0$.*
2) *La funzione f è continua in x_0 se per ogni $j \in \{1, \cdots, p\}$ f_j è continua in x_0.*

3. Funzioni differenziabili

Diamo la definizione per una funzione $f : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$ e poi quella generale. Poiché si tratta di una proprietà locale formuliamo la definizione e i seguenti risultati supponendo che Ω sia aperto.

Definizione 3.1. *Sia $f : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$, $x_0 \in int(\Omega) = \Omega$. Diremo che f è differenziabile in x_0 se esiste $T \in L(\mathbb{R}^2, \mathbb{R})$ ($L(\mathbb{R}^2, \mathbb{R})$ è l'insieme delle applicazioni lineari da $\mathbb{R}^2$ a $\mathbb{R}$), applicazione lineare, tale che*

$$f(x_0 + h) - f(x_0) = T(h) + o(||h||), \ \ h \to 0.$$

In tal caso chiameremo T un differenziale di f in x_0.

Teorema 3.1. *Sia $f : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$, $x_0 \in int(\Omega) = \Omega$. f è differenziabile se e solo se esiste $T \in L(\mathbb{R}^2, \mathbb{R})$, applicazione lineare, tale che*

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0) - T(h)}{||h||} = 0.$$

Vediamo nel caso di una funzione ad una variabile che cosa significa affermare che $f : I \subset \mathbb{R} \to \mathbb{R}$, con I intervallo, è differenziabile in $x_0 \in \mathrm{int}(I)$. Supponiamo per semplicità che f sia derivabile in x_0. Allora

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} = f'(x_0) \in \mathbb{R}.$$

Poniamo $h \to \omega(x_0, h) := \frac{f(x_0+h)-f(x_0)}{t} - f'(x_0)$ per ogni x in un intorno di x_0, $x_0 \neq 0$. Quindi risulta

$$f(x_0 + h) - f(x_0) - f'(x_0)h = \omega(x_0, h)h.$$

Pertanto l'applicazione $df(x_0) : \mathbb{R} \to \mathbb{R}$

$$h \to f'(x_0)h$$

è lineare ed è il differenziale di f in x_0, perché $\omega(x_0, h)h = o(h)$ per $h \to 0$.

Analogamente se sappiamo che $f : I \to \mathbb{R}$ è differenziabile in x_0, allora esiste un'applicazione lineare, indichiamola ancora con $df(x_0) : \mathbb{R} \to \mathbb{R}$ tale che

$$f(x_0 + h) - f(x_0) = df(x_0)(h) + o(h).$$

D'altra parte un'applicazione lineare da $\mathbb{R}$ in $\mathbb{R}$ è univocamente determinata dal valore assunto dall'applicazione stessa dal vettore della base, in questo caso 1. Infatti per le proprietà di linearità di $df(x_0)$ risulta $df(x_0)(h) = hdf(x_0)(1)$ con $df(x_0)(1) \in \mathbb{R}$. Pertanto

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} = df(x_0)(1) \in \mathbb{R},$$

cioè f è derivabile in x_0 e la derivata di f in x_0 è $df(x_0)(1)$. Quindi per funzioni reali a valori reali risulta che f è differenziabile in un punto se e solo se in quel punto è derivabile. In più variabili non vale più questa equivalenza. Nei prossimi risultati forniremo alcune ulteriori precisazioni senza addentrarci ulteriormente nella questione perché esula dagli scopi di questo corso.

Teorema 3.2. *Sia $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$, per semplicità possiamo supporre $n = 2$ con Ω aperto e $x_0 \in \Omega$. Se f è differenziabile in x_0, allora*

i) *f è continua in x_0.*
ii) *f ha derivata rispetto ad ogni vettore $a\mathbb{R}^n$, $||a|| = 1$ e se T è un differenziale di f in x_0, allora*
$$T(a) = \frac{\partial f(x_0)}{\partial a}.$$
iii) *il differenziale è unico e lo indicheremo con il simbolo*
$$df(x_0) : \mathbb{R}^n \to \mathbb{R}$$
la matrice, in questo caso il vettore, che rappresenta l'applicazione differenziale è detta matrice Jacobiana di f in x_0 ($Jf(x_0)$), nel caso specifico di una sola riga (cioè in questo caso) si parla di gradiente di f in x_0 il cui simbolo è $\nabla f(x_0)$.

Ricordiamo che ogni applicazione lineare è univocamente determinata dai valori che assume sui vettori della base, quindi se f è differenziabile in x_0, allora

$$\begin{aligned} df(x_0)(h) &= df(x_0)(\sum_{i=1}^n h_i e_i) = \sum_{i=1}^n h_i df(x_0)(e_i) \\ &= \sum_{i=1}^n h_i \frac{\partial f(x_0)}{\partial x_i} = \langle \nabla f(x_0), h \rangle. \end{aligned}$$

Definizione 3.2. *(Definizione di derivata parziale) Sia $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$, (per semplicità possiamo supporre $n = 2$). Diremo che f ha derivata parziale rispetto a x_j, $j \in \{1, \cdots, n\}$ se esiste*

$$\lim_{t \to 0} \frac{f(x_0 + te_j) - f(x_0)}{t} = \alpha_j \in \mathbb{R},$$

dove $\{e_1, \cdots, e_n\}$ è la base canonica di $\mathbb{R}^n$. Tale limite α_j è detto la derivata parziale di f rispetto a x_j calcolata in x_0 e viene indicato con il seguente simbolo

$$\frac{\partial f(x_0)}{\partial x_j}.$$

A livello di notazione diremo che $f : \Omega \subset \mathbb{R}^n :\to \mathbb{R}$, Ω aperto, è di classe e scriveremo $f \in C^1(\Omega)$ se per ogni punto x_0 f è dotata di tutte le derivate parziali e queste derivate sono continue in ogni punto di Ω.

Per completezza diamo la definizione di funzione vettoriale differenziabile.

Definizione 3.3. *Sia $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^p$, Ω aperto e $x_0 \in \Omega$. Diremo che f è differenziabile in x_0 se per ogni componente f_j, $j = 1, \ldots, p$, $f = (f_1, \ldots, f_p)$ f_j è differenziabile in x_0. Chiameremo differenziale di f in x_0 l'applicazione lineare $df(x_0) \in L(\mathbb{R}^n, \mathbb{R}^p)$, ($L(\mathbb{R}^n, \mathbb{R}^p)$ indica l'insieme di tutte l' applicazioni lineari da $\mathbb{R}^n$ a $\mathbb{R}^p$) così definita*

$$df(x_0)(h) = (df_1(x_0)(h), \ldots, df_p(x_0)(h)),$$

dove $df_1(x_0), \ldots, df_p(x_0)$ sono i differenziali delle componenti $f_1, \ldots, f_p$ in x_0. La matrice che rappresenta il differenziale di f in x_0 è una matrice di p righe e n colonne detta matrice Jacobiana e che prende la seguente forma:

$$Jf(x_0) = \begin{bmatrix} \frac{\partial f_1(x_0)}{\partial x_1}, & \frac{\partial f_1(x_0)}{\partial x_2}, & \ldots, & \frac{\partial f_1(x_0)}{\partial x_n} \\ \frac{\partial f_2(x_0)}{\partial x_1}, & \frac{\partial f_2(x_0)}{\partial x_2}, & \ldots, & \frac{\partial f_2(x_0)}{\partial x_n} \\ \ldots, & \ldots, & \ldots, & \ldots \\ \frac{\partial f_p(x_0)}{\partial x_1}, & \frac{\partial f_p(x_0)}{\partial x_2}, & \ldots, & \frac{\partial f_p(x_0)}{\partial x_n}. \end{bmatrix}$$

Vale il segunete risultato che fornisce una condizione sufficiente affinché f sia differenziabile e costituisce una parziale motivazione a quanto sottolineato in merito nel caso di funzioni da valori reali a valori reali.

Teorema 3.3. *Sia $\Omega \subset \mathbb{R}^n$, $f \in C^1(\Omega)$, Ω aperto. Allora f è differenziabile in ogni punto di Ω.*

Per $n > 1$ non vale il viceversa, cioè una funzione può essere differenziabile in un punto senza che sia di classe C^1 in alcun intorno di x_0.

Teorema 3.4. *Sia $\Omega \subset \mathbb{R}^n$, $f, g : \Omega \to \mathbb{R}^p$, Ω aperto e $x_0 \in \Omega$.*

i) *Se f, g sono differenziabili in x_0, allora $f + g$ è differenziabile in x_0 e*

$$d(f+g)(x_0) = df(x_0) + dg(x_0).$$

ii) *Per ogni $\alpha \in \mathbb{R}$ cf è differenziabile in x_0 e*

$$d(cf)(x_0) = cdf(x_0).$$

Teorema 3.5. *Sia $\Omega \subset \mathbb{R}^n$, $f : \Omega \to \mathbb{R}^p$, Ω aperto e f sia differenziabile in $x_0 \in \Omega$. Supponiamo inoltre che $f(x_0) \in int(f(\Omega)) \subset B$ con $g : B \subset \mathbb{R}^p \to \mathbb{R}^q$ e g differenziablie in $f(x_0)$. Allora $g \circ f : \Omega \to \mathbb{R}^q$ è differenziabile in x_0 e*

$$d(f \circ g)(x_0) = dg(f(x_0)) \circ df(x_0).$$

Inoltre

$$J(f \circ g)(x_0) = Jg(f(x_0)) \cdot Jf(x_0).$$

In particolare l'entrata ik-esima della precedente matrice corrisponde al prodotto scalare della i-esima riga della matrice $Jg(f(x_0))$ per la k-esima colonna della matrice $Jf(x_0)$ ovvero

$$\frac{\partial (g\circ f)_i(x_0)}{\partial x_k} = [\frac{\partial g_i(f(x_0))}{\partial y_1}, \frac{\partial g_i(f(x_0))}{\partial y_2}, \dots, \frac{\partial g_i(f(x_0))}{\partial y_p}] \cdot \begin{bmatrix} \frac{\partial f_1(x_0)}{\partial x_k} \\ \frac{\partial f_2(x_0)}{\partial x_k} \\ \cdots \\ \frac{\partial f_p(x_0)}{\partial x_k} \end{bmatrix} = \sum_{m=1}^{p} \frac{\partial g_i(f(x_0))}{\partial y_m} \frac{\partial f_m(x_0)}{\partial x_k}$$

Vediamo alcune applicazioni del precedente risultato. Sia $f\mathbb{R}^3 \to \mathbb{R}$ e $\phi : I \subset \mathbb{R} \to \mathbb{R}^3$, I intervallo di $\mathbb{R}$. Supponiamo che entrambe le funzioni siano differenziabili. Allora

$$J(f \circ \phi)(t_0) = Jf(\phi(t_0)) \cdot J\phi(t_0) = \langle \nabla f(\phi(t_0)), \phi'(t_0) \rangle,$$

dove $\phi'(t_0) = (\phi_1'(t_0), \phi_2'(t_0), \phi_3'(t_0))$.

Calcolare il gradiente nel punto (x_0, y_0) della seguente funzione $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = g(x\exp(x^2y), h(x,y))$ dove $h, g : \mathbb{R}^2 \to \mathbb{R}$ sono differenziabili in $\mathbb{R}^2$.

$$\begin{aligned} \frac{\partial f(x,y)}{\partial x} &= \frac{\partial g(x_0\exp(x_0^2y_0), h(x_0,y_0))}{\partial u_1}(\exp(x_0^2y_0) + 2x_0^2y_0\exp(x_0^2y_0)) \\ &+ \frac{\partial g(x_0\exp(x_0^2y_0), h(x_0,y_0))}{\partial u_2}\frac{\partial h(x_0,y_0)}{\partial x}. \\ \frac{\partial f(x,y)}{\partial y} &= \frac{\partial g(x_0\exp(x_0^2y_0), h(x_0,y_0))}{\partial u_1}x_0^3\exp(x_0^2y_0) \\ &+ \frac{\partial g(x_0\exp(x_0^2y_0), h(x_0,y_0))}{\partial u_2}\frac{\partial h(x_0,y_0)}{\partial y}. \end{aligned}$$

Scrivere la matrice Jacobiana della seguente funzione $f : \mathbb{R}^2 \to \mathbb{R}^3$,

$$f(x,y) = (\sin(xy^4), 2 - x + y, y\log(1+x^2))$$

$$Jf(x_0) = \begin{bmatrix} y^4\cos(xy^4), & 4y^3\cos(xy^4) \\ -1, & 1 \\ \frac{2xy}{1+x^2}, & \log(1+x^2). \end{bmatrix}$$