Richiami di Analisi II

Kevin R. Payne
Appunti per il Corso Analisi Matematica III
CdL in Matematica e Matematica per le Applicazioni
Università degli Studi di Milano

settembre 2004

In questi appunti vogliamo ricordare brevemente le notazioni, concetti di base, risultati di base che vengono forniti nel corsi di Analisi Matematica II. Sono consigliati gli appunti di Analisi II dei Proff. Salvatori e Vignati [1] per un trattemento più esteso ma ancora snello.

1 Lo spazio $\mathbf{R}^n$

1.1 Strutture lineare, geometriche, topologiche

Denotiamo con $\mathbf{R}^n$ lo *spazio euclideo* di dimensione n; cioè

$$\mathbf{R}^n = \{x = (x_1, \ldots, x_n) : x_k \in \mathbf{R}, \quad k = 1, \ldots, n\} \tag{1.1}$$

con le due operazioni di *addizione*

$$x + y = (x_1 + y_1, \ldots, x_n + y_n), \quad x, y \in \mathbf{R}^n \tag{1.2}$$

e *molteplicazione scalare*

$$\lambda x = (\lambda x_1, \ldots, \lambda x_n) \tag{1.3}$$

Denotiamo con $\{e_j\}_{j=1}^n$ la base canonica di $\mathbf{R}^n$ dove $[e_j]_k = \delta_{jk}$ vale 1 se $j = k$ e 0 per $j \neq k$. Notiamo che spesso un elemento generico di $\mathbf{R}^n$ viene sritto come $\vec{x}$ oppure come $\underline{x}$ per dare enfasi sul fatto che $\mathbf{R}^n$ è uno spazio vettoriale (di dimensione n); qui invece seguiamo la strada di notarlo solo con x, notazione più pulita in cui è necessario solo fare attenzione ai dichiarazioni da dove "vive" x.

Su $\mathbf{R}^n$ abbiamo bisogno di due applicazioni che forniscono la base delle proprietà topologiche e geometriche; cioè ricordiamo che il *prodotto scalare* $\langle \cdot, \cdot \rangle : \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$ è definita da

$$\langle x, y \rangle = \sum_{k=1}^{n} x_k y_k \tag{1.4}$$

e la *norma euclidea* $|| \cdot || : \mathbf{R}^n \to [0, \infty) \subset \mathbf{R}$ è definita da

$$||x|| = \langle x, x \rangle^{1/2} = \sqrt{x_1^2 + \ldots + x_n^2} \tag{1.5}$$

Il prodotto scalare è un'applicazione bilineare, simmetrica e nondegenere (si ricordi cosa significa) e quindi la norma è anche nondegenere ($||x|| = 0$ se e solo se $x = 0$) che è legata alla struttura lineare su $\mathbf{R}^n$ via

$$||\lambda x|| = |\lambda| \, ||x||, \quad \lambda \in \mathbf{R}, x \in \mathbf{R}^n \tag{1.6}$$

$$||x + y|| \leq ||x|| + ||y||, \quad x, y \in R^n \tag{1.7}$$

Osservazione 1: *Queste funzioni definiscono su $\mathbf{R}^n$ le proprietà topologiche (insiemi aperti, chiusi, compatti, funzioni/applicazioni continue) e le proprietà geometriche (lunghezza, distanza, angoli).*

Osservazione 2: *Queste due funzioni $\langle \cdot, \cdot \rangle$ e $|| \cdot ||$ hanno valori in $\mathbf{R}$ (un campo ordinato completo) e sono collegate tramite la disuguaglianza di Cauchy-Schwarz*

$$|\langle x, y \rangle| \leq ||x|| \, ||y||, \quad x, y \in \mathbf{R}^n \tag{1.8}$$

In particolare, dalla norma (oppure dal prodotto scalare) si definisce una *funzione di distanza (metrica)* $d : \mathbf{R}^n \times \mathbf{R}^n \to [0, \infty)$ come

$$d(x, y) = ||x - y|| = \langle x - y, x - y \rangle^{1/2} = \left(\sum_{k=1}^{n} (x_k - y_k)^2 \right)^{1/2} \tag{1.9}$$

La distanza è una funzione simmetrica, nondegenere, che soddisfa la disuguaglianza tringolare

$$d(x,y) \leq d(x,z) + d(z,y), \quad x,y,z \in \mathbf{R}^n \tag{1.10}$$

Così $(\mathbf{R}^n, d)$ diventa uno *spazio metrico.*

1.2 Sottoinsiemi particolari

Usiamo spesso sottoinsiemi di $\mathbf{R}^n$ di semplice geometria e quindi forniamo una notazione per loro. Abbiamo la *palla aperta (intorno circolare aperto) di raggio* $r > 0$ *e centro* x_0,

$$B_r(x_0) = \{x \in \mathbf{R}^n : d(x,x_0) = ||x - x_0|| < r\} \tag{1.11}$$

la *palla chiusa di raggio* r *e centro* x_0,

$$\overline{B}_r(x_0) = \{x \in \mathbf{R}^n : d(x,x_0) = ||x - x_0|| \leq r\} \tag{1.12}$$

la sfera di raggio r con centro x_0

$$S_r(x_0) = \{x \in \mathbf{R}^n : d(x,x_0) = ||x - x_0|| = r\} \tag{1.13}$$

che è la *frontiera* anche detto *bordo* di $B_r(x_0)$, ed il *segmento fra* x *e* y *in* R^n che è indicato con

$$[x,y] = \{z \in \mathbf{R}^n : z = (1-t)x + ty, \quad 0 \leq t \leq 1\}. \tag{1.14}$$

1.3 Concetti topologici/geometrici

Ricordiamo che un insieme $A \subseteq \mathbf{R}^n$ si chiama *aperto* se e solo se $\forall x_0 \in A \; \exists r = r(x_0)$ tale che $B_r(x_0) \subset A$ e che A è chiuso se e solo se il suo complemento $A^C = \mathbf{R}^n \setminus A$ sia aperto. Dato un'insieme $A \subset \mathbf{R}^n$ denotiamo con $\overline{A}$ la *chiusura di* A *in* $\mathbf{R}^n$, ∂A la *frontiera (il bordo) di* A, $A^\circ = \overline{A} \setminus \partial A$ l'interno di A.

Esercizio 1: Mostrare che $B_r(x_0)$ è aperto in $\mathbf{R}^n$ e che $\overline{B}_r(x_0), S_r(x_0), [x,y]$ sono chiusi in $\mathbf{R}^n$.

Ricordiamo che $A \subset \mathbf{R}^n$ è *limitato* se e solo se $\exists x_0 \in \mathbf{R}^n$, $\exists R > 0$ tali che $A \subseteq B_R(x_0)$. Possiamo usare la funzione di distanza per caraterizzare la limitatezza di un insieme e di fornire una definizione di *diametro* per un insieme limitato; cioè

$$\operatorname{diam}(A) = \sup\{d(x,y) : x, y \in A\}, \quad A \text{ limitato, non vuoto} \tag{1.15}$$

Proponiamo qualche esercizio.

Esercizio 2: Mostrare che un insieme è limitato se e solo se la distanza su A, $d_{|A} : A \times A \to [0, +\infty)$ è limitata (superiormente). Mostrare che il diametro è una funzione crescente ripsetto la relazione parziale di inclusione sugli insiemi limitati.

Esercizio 3: Verificare che $\operatorname{diam}(B_r(x_0)) = 2r$. Calcolare il diametro per gli insiemi

$$Q_r(0) := \{x \in \mathbf{R}^n : \quad ||x||_\infty := \max_{1 \le i \le n} |x_i| < r\}$$
$$\mathcal{D}_r(0) := \{x \in \mathbf{R}^n : \quad ||x||_1 := \sum_{i=1}^{n} |x_i| < r\}$$

Esercizio 4: Sia $A \subset \mathbf{R}^n$ limitato, non vuoto. Mostrare che è ben definito il seguente *diametro esterno*

$$\operatorname{diam}_e(A) := 2 \inf\{R \in (0, \infty) : \exists x_0 \in A \ \text{ t.c. } \ A \subseteq B_R(x_0)\} < +\infty.$$

Nel caso di $A = B_r(x_0)$ abbiamo $\operatorname{diam}(A) = \operatorname{diam}_e(A)$. In generale c''e una relazione fra diam e diam_e?

La topologia di $\mathbf{R}^n$ viene determinata dalla sua base di intorni circolari aperti che fornisce i concetti essenziali di convergenza e compattezza. In particolare ricordiamo che *una successione di punti* $\{x_j\}_{j=1}^{\infty} \subset \mathbf{R}^n$ *converge ad* $x \in \mathbf{R}^n$ se e solo se definitivamente $x_j \in B_\epsilon(x)$ per ogni $\epsilon > 0$. Inoltre, essendo che questi intorni sono definiti tramite la norma (oppure la metrica) che ha valori in $\mathbf{R}$, abbiamo

$$
\begin{aligned}
x_j \to x \ \text{ in } \mathbf{R}^n \quad &\Leftrightarrow \quad \forall\, \epsilon > 0\ \exists\ N = N(\epsilon) \text{ t.c. } ||x_j - x|| < \epsilon, \forall j > N \\
&\Leftrightarrow \quad \lim_{j\to+\infty} ||x_j - x|| = 0 \\
&\Leftrightarrow \quad \lim_{j\to+\infty} d(x_j, x) = 0
\end{aligned} \tag{1.16}
$$

Osservazione 3: $(\mathbf{R}^n, d)$ *è uno spazio metrico* completo*; cioè* $\{x_j\}$ *converge ad* $x \in \mathbf{R}^n$ *se e solo se la successione è una successione di Cauchy*

$$
\forall \epsilon > 0 \ \ \exists \ \ N = N(\epsilon) \ \text{ t.c. } \ ||x_j - x_k|| < \epsilon \ \ \forall j, k \geq N \tag{1.17}
$$

Ricordiamo che un sottoinsieme A di $\mathbf{R}^n$ si chiama *compatto* se e solo se dato un ricoprimento (non necessariamente numerabile) di A tramite intorni aperti esiste una sottocollezione finita che rimane un ricoprimento; cioè

$$
A \subseteq \cup_{j\in J} B_{r_j}(x_j) \Rightarrow \exists j_1, \dots j_N \ \text{ t.c. } \ A \subseteq B_{r_{j_1}}(x_{j_1}) \cup \cdots \cup B_{r_{j_N}}(x_{j_N}). \tag{1.18}
$$

Più facilmente si sa il sequente caratterizzazione dei compatti in $\mathbf{R}^n$.

Osservazione 4: (Teorema di Heine-Borel) $A \subset \mathbf{R}^n$ *è compatto se e solo se* A *è chiuso e limitato.*

Una conseguenza del Teorema di Bolzano-Weierstrass è il seguente uso commune della compatezza in $\mathbf{R}^n$.

Osservazione 5: *Sia* $\{x_j\} \subset K$ *dove* K *è compatto. Allora esiste una sottosuccessione* $\{x_{j_k}\}$ *convergente ad un punto* $x \in K$.

Finalmente ricordiamo che un sottoinsieme aperto A di $\mathbf{R}^n$ si chiama *connesso* se e solo se A **non** può essere scritto come l'unione di due insiemi aperti, disgiunti e non vuoti.

Osservazione 6: *Un insieme aperto* $A \subseteq \mathbf{R}^n$ *è connesso se e solo se è connesso per archi; cioè per ogni* $x, y \in A$ *esiste un cammino continuo che fra* x *ad* y *con sostegno in* A.

Vedi pargafrafo 4.2.3 di [1] per una variante via *poligonali* in cui non è necessario di sapere già cosa significa *cammino continuo* ($\varphi : [0, 1] \to A$ continua tale che $\varphi(0) = x$ e $\varphi(1) = y$).

2 Limiti e continuità per funzioni di più variabili

Ricordiamo i concetti di limiti e continuità per funzioni a valori vettoriali in più variabili

$$\begin{aligned} f &: A \subseteq \mathbf{R}^n \to \mathbf{R}^m \\ x &\mapsto f(x) = (f_1(x), \ldots, f_m(x)) \end{aligned}$$

dove spesso assumiamo che l'interno di A non sia vuoto per cui esistano punti nel dominio non isolati; in generale la questione di limiti/continuità è interessante solo vicino un *punto di accumulazione* x_0 nel dominio. Le norme euclidee nel dominio e codominio definiscono queste proprietà.

2.1 Limiti

Ricordiamo che f *ha limite* $l \in \mathbf{R}^m$ e scriviamo $\lim_{x \to x_0} f(x) = l$ se e solo se

$$\forall \ \ \epsilon > 0 \ \ \exists \ \ \delta = \delta(\epsilon): \ \ [x \in A] \wedge [0 < ||x - x_0||_{\mathbf{R}^n} < \delta] \Rightarrow ||f(x) - l||_{\mathbf{R}^m} < \epsilon \tag{2.1}$$

dove si possono sempre redurre la questione di convergenza di f a quella delle componenti f_k di f; cioè si sa

$$f(x) \to l \Leftrightarrow f_k(x) \to l_k \ \ \forall \ k = 1, \ldots m \tag{2.2}$$

e quindi si possono sfruttare risultati/strumenti unidimensionali. Ricordiamo inoltre due criteri per il calcolo di limiti che spesso vengono usati per mostrare la nonesistenza di limiti.

Osservazione 7: $\lim_{x \to x_0} f(x) = l$ *se e solo se il limite di* f *ristretta lungo ogni cammino continuo con sostegno in* A *che temina in* x_0 *esiste e vale* l.

Osservazione 8: *Nel piano,* $\lim_{(x,y) \to (x_0,y_0)} f(x) = l$ *se e solo se* $f(x_0 + r\cos\theta, y_0 + r\sin\theta) \to l$ *quando* $r \to 0^+$ *uniformemente in* θ.

Per qualche esercizio su questi criteri si può consultare [1] (paragrafo 4.3.3).

2.2 Continuità

Ricordiamo infine che una funzione f è *continua in* $x_0 \in A^\circ$ se e solo se

$$\lim_{x \to x_0} f(x) = f(x_0) \tag{2.3}$$

dove di nuovo nel caso di valori in $\mathbf{R}^m$ basta controllare la continuità dei componenti di f. Si dice f *continua su* A quando f è continua in ogni punto x_0 di A. Valgono i soliti risultati di continuità per la somma, composizione, etc. di funzioni continue. Inoltre **solo** nel caso di funzioni a valori in $\mathbf{R}$ (dove c'è una struttura di ordine) abbiamo generalizzazioni dei teoremi dei valori intermedi e di Weierstrass.

Osservazione 9: (Teorema dei valori intermedi) *Sia* $f : A \subset \mathbf{R}^n \to \mathbf{R}$ *dove* A *è aperto e connesso e* f *è continua su* A *con* $f(x) = a < b = f(y)$ *per* $x, y \in A$. *Allora per ogni* $c \in (a, b)$ *esiste* $z \in A$ *tale che* $f(z) = c$.

Osservazione 10: (Teorema di Weierstrass) *Sia* $f : K \subset \mathbf{R}^n \to \mathbf{R}$ *continua su* K *compatto. Allora esistono* $x_m, x_M \in K$ *tali che*

$$f(x_m) = m = \min_{x \in K} f(x) \quad \text{e} \quad f(x_M) = M = \max_{x \in K} f(x) \tag{2.4}$$

Proponiamo qualche esercizio semplice.

Esercizio 5: Verificare che la norma $|| \cdot || : \mathbf{R}^n \to \mathbf{R}$ è continua.

Esercizio 6: Verificare che la proiezione

$$\pi_k : \mathbf{R}^n \to \mathbf{R}$$

$$x = (x_1, \ldots, x_n) \mapsto x_k$$

è continua per ogni $k = 1, \ldots, n$.

Esercizio 7: Mostrare che esistono due costanti positivi C_1, C_2 (independenti da $x \in \mathbf{R}^n$) tali che

$$C_1 ||x|| \leq \max_{1 \leq k \leq n} |x_k| \leq C_2 ||x||. \tag{2.5}$$

Quest'ultimo esercizio è utile per verificare che la continuità di una funzione a valori in $\mathbf{R}^m$ equivale la continuità delle componenti. Inoltre dice che la norma $||\cdot||$ è *equivalente* alla norma $|||x||| = \max_{1\leq k\leq n}|x_k|$. Infatti, si può mostrare che ogni copia di norme su $\mathbf{R}^n$ sono equivalenti nel senso della formula (2.5), perchè?

3 Calcolo differenziabile per funzioni di più variabili

3.1 Derivate direzionali e derivate parziali

Ricordiamo che data $f : A \subset \mathbf{R}^n \to \mathbf{R}^m$ con $x_0 \in A$ e v un versore di $\mathbf{R}^n$ ($||v|| = 1$) la *derivata direzionale di f nella direzione v nel punto x_0* è definita da

$$\frac{\partial f}{\partial v}(x_0) = \lim_{t\to 0}\frac{1}{t}[f(x_0+tv) - f(x_0)] \tag{3.1}$$

se esiste finito questo limite e diciamo che f *è derivabile nella direzione v nel punto x_0*. Nel caso speciale di $v = e_j$ parliamo anche della *derivata parziale di f rispetto ad x_j nel punto x_0*; cioè

$$\frac{\partial f}{\partial x_j}(x_0) = \lim_{t\to 0}[f(x_0+te_j) - f(x_0)] \tag{3.2}$$

Quando f ha valori vettoriali queste derivate possono essere calcolate componente per componente; cioè

$$\frac{\partial}{\partial v}\begin{bmatrix} f_1 \\ \vdots \\ f_m \end{bmatrix} = \begin{bmatrix} \partial f_1/\partial v \\ \vdots \\ \partial f_m/\partial v \end{bmatrix} \tag{3.3}$$

Quando esistano tutte le derivate parziali di f possiamo formare il *gradiente di f*; cioè

$$\nabla f(x_0) = ((\partial f/\partial x_1)(x_0), \ldots, (\partial f/\partial x_n)(x_0)) \tag{3.4}$$

che è un vettore in $\mathbf{R}^n$ nel caso di f scalare. Spesso usiamo le notazioni $D_v f$ per $\partial f/\partial v$ e $D_{x_j}f = f_{x_j}$ per $\partial f/\partial x_j$. Ricordiamo che le *derivate parziali di ordine superiore* sono definite in modo naturale

$$D^2_{x_k x_j} f = D_{x_k}(D_{x_j} f) = \frac{\partial}{\partial x_k}\left(\frac{\partial f}{\partial x_j}\right) = \frac{\partial^2 f}{\partial x_k \partial x_j} = \left(f_{x_j}\right)_{x_k} = f_{x_j x_k} \tag{3.5}$$

e così via. Introduciamo la notazione di *multi-indici* per le derivate parziali; cioè dato $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbf{N}^n$ si denota

$$D^\alpha_x f = D^{\alpha_1}_{x_1} \cdots D^{\alpha_n}_{x_n} f \tag{3.6}$$

dove $|\alpha| = \alpha_1 + \ldots \alpha_n$ è l'ordine della derivata parziale.

Il concetto di derivate parziale è alle base delle leggi che descrivono modelli per fenomeni della scienza; in particolare si tratta di *equazioni alle derivate parziali* che sono anche di interesse nella geometria e topologia differenziali. Proponiamo qualche esercizio per iniziare un discorso che si ripresenterà spesso in Analisi III.

Esercizio 8: Siano $f, g \in C^2(\mathbf{R})$ (due volte derivabile con continuità). Mostrare che la funzione

$$u(x,t) = f(x+t) + g(x-t)$$

è *soluzione dell'equazione delle onde (D'Alembert 1746)*; cioè

$$\frac{\partial^2 u}{\partial^2 t}(x,t) - \frac{\partial^2 u}{\partial^2 x}(x,t) = 0, \quad (x,t) \in \mathbf{R}^2 \tag{3.7}$$

In realtà tutte le funzioni $u(x,t)$ di classe $C^2(\mathbf{R}^2)$ che soddisfano l'equazione (3.7) sono di questa forma. L'equazione delle onde è un buon modello per le piccole vibrazioni di una corda elastica (violino, chitarra, etc.) e generalizzazioni di quest'equazione sono modelli di quantità dinamiche che si propagano come "onde" (luce, suono, etc.)

Esercizio 9: Mostrare che la funzione $u(x_1, x_2) = \ln\left(\sqrt{x^2 + x_2^2}\right)$ è *soluzione dell'equazione di Laplace (1840)*; cioè

$$\frac{\partial^2 u}{\partial^2 x_1}(x_1,x_2) + \frac{\partial^2 u}{\partial^2 x_2}(x_1,x_2) = 0, \quad (x_1,x_2) \in \mathbf{R}^2 \setminus (0,0) \tag{3.8}$$

Le soluzioni dell'equazione di Laplace sono dette *funzioni armoniche* e sono le protagoniste del corso avvanzato. Loro giocano un ruolo fondamentale nella scienza, nella geometria e nel analisi complessa.

3.2 Differenziabilità

Ricordiamo che per funzioni di più di una variabile l'esistenza delle derivate direzionali ed in particolare le derivate parziali **non** implica in generale la continuità della funzione e quindi abbiamo bisogno di un concetto più forte. Data $f : A \subseteq \mathbf{R}^n \to \mathbf{R}^m$ e $x_0 \in A^\circ$ si dice f *è differenziabile in* x_0 se e solo se esiste un'applicazione lineare $L : \mathbf{R}^n \to \mathbf{R}^m$ tale che

$$\lim_{h \to 0} \frac{1}{||h||_{\mathbf{R}^n}} [f(x_0 + h) - f(x_0) - Lh] = 0 \tag{3.9}$$

dove $h \to 0 \in \mathbf{R}^n$ equivale $||h||_{\mathbf{R}^n} \to 0$ ed il limite è preso nel senso di convergenza in $\mathbf{R}^m$. In tal caso, si chiama $L = df(x_0)$ il *differenziale di* f *in* x_0. Dall'algebra lineare sappiamo che a quest'applicazione lineare da $\mathbf{R}^n$ in $\mathbf{R}^m$ insieme con una scelta di base per $\mathbf{R}^n$ e $\mathbf{R}^m$ si può associare una matrice $Df(x_0)$ tale che

$$df(x_0)h = [Df(x_0)][h] \tag{3.10}$$

Con la scelta delle base canoniche la matrice $Df(x_0)$ si chiama *matrice jacobiana di* $f = (f_1, \ldots, f_m)$ *in* x_0 e si mostra che

$$Df(x_0) = \left[\begin{array}{ccc} \partial f_1/\partial x_1 & \cdots & \partial f_1/\partial x_n \\ \vdots & \ddots & \vdots \\ \partial f_m/\partial x_1 & \cdots & \partial f_m/\partial x_n \end{array} \right]_{x=x_0} = \left[\begin{array}{ccc} \longleftarrow & \nabla f_1(x_0) & \longrightarrow \\ & \vdots & \\ \longleftarrow & \nabla f_m(x_0) & \longrightarrow \end{array} \right] \tag{3.11}$$

Come si nota dalla notazione, nella matrice jacobiana le righe sono i gradienti delle componenti di f e le colonne sono le derivate parziali di f. Il significato della differenziablitità è che valori $f(x_0+h)$ della funzione f sono *ben approssimate* dalla funzione *affine* $f(x_0)+Lh$ per h abbastanza piccolo, come accade nel caso di funzioni derivabili da $\mathbf{R}$ in $\mathbf{R}$.

Osservazione 11: *Sia* $f : A \subset \mathbf{R}^n \to \mathbf{R}^m$ *differenziabile in* x_0*. Allora*

$$f(x_0 + h) = f(x_0) + df(x_0)h + R_1(x_0, h) \tag{3.12}$$

dove

$$||R_1(x_0, h)||_{\mathbf{R}^m} = o(||h||_{\mathbf{R}^n}) \;\text{ per }\; h \to 0 \;\text{ in }\; \mathbf{R}^n. \tag{3.13}$$

Ricordiamo alcuni spazi di funzioni con A un aperto di $\mathbf{R}^n$

$$C^0(A, \mathbf{R}^m) = \{f : A \to \mathbf{R}^m : f \text{ sia continua in } A\} \tag{3.14}$$

$$\mathrm{Diff}(A, \mathbf{R}^m) = \{f : A \to \mathbf{R}^m : f \text{ sia differenziabile in } A\} \tag{3.15}$$

$$C^k(A, \mathbf{R}^m) = \{f : A \to \mathbf{R}^m : D_x^\alpha f \in C^0(A, \mathbf{R}^m), \quad \forall |\alpha| \leq k\}. \tag{3.16}$$

Spesso scriviamo solo $C^0(A)$ etc. quando sia chiaro il contesto (soprattuto nel caso di funzioni scalari ($m = 1$)). In ogni caso, questi spazi formano una catena ripetto l'inclusione

$$C^k(A) \subset C^{k-1}(A) \subset \cdots \subset C^1(A) \subset \mathrm{Diff}(A) \subset C^0(A) \tag{3.17}$$

e poniamo

$$C^\infty(A) = \cap_{k \in N} C^k(A). \tag{3.18}$$

Per fare dei conti nel calcolo differenziabile dobbiamo ricordare come derivare prodotti, somme, e sopratutto composizioni di funzioni. Ricordiamo esplicitamente questo risultato fondamentale.

Osservazione 12: (la regola della catena) *Siano $f : A \subset \mathbf{R}^n \to \mathbf{R}^m$ e $g : B \subset \mathbf{R}^m \to \mathbf{R}^l$ differenziabili con A, B aperti e $f(A) \subseteq B$. Allora la funzione composta $g \circ f$ è differenziabile in A e vale*

$$D(g \circ f)(x) = Dg(f(x))Df(x), \quad x \in A \tag{3.19}$$

3.3 La formula di Taylor

L'osservazione 11 è nient'altro di una formula di Taylor di ordine 1 con resto di Peano per f basata in x_0 che equivale (mettendo $h = x - x_0$)

$$f(x) = f(x_0) + df(x_0)(x - x_0) + R_1(x_0, h) \tag{3.20}$$

dove

$$||R_1(x_0, h)||_{\mathbf{R}^m} = o(||h||_{\mathbf{R}^n}) \text{ per } h \to 0 \text{ in } \mathbf{R}^n. \tag{3.21}$$

Nel caso scalare con $f : A \subset \mathbf{R}^n \to \mathbf{R}$ con f differenziabile in x_0 la formula di Taylor prende la forma

$$\begin{aligned} f(x) &= f(x_0) + \langle \nabla f(x_0), x - x_0 \rangle + o(||x - x_0||) \quad \text{per} \quad x \to x_0 \in \mathbf{R}^n && (3.22) \\ &= f(x_0) + \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(x_0)(x - x_0) + o(||x - x_0||) \quad \text{per} \quad x \to x_0 \in \mathbf{R}^n && (3.23) \end{aligned}$$

Si osserva che la funzione $g(x) = f(x_0) + \langle \nabla f(x_0), x - x_0 \rangle$ ha come grafico un iperpiano che passa per il punto $(x_0, f(x_0)) \in \mathbf{R}^n \times \mathbf{R}$; questo piano si chiama *piano (spazio) tangente al grafico di f nel punto* $(x_0, f(x_0))$.

Esercizio 10: Trovare l'equazione del piano tangente al grafico di $f(x_1, x_2) = x_1^2 + x_2^2$ nei punti $x_0 = (0, 0)$ e $x_0 = (1, 0)$.

Per una funzione $f : A \subset \mathbf{R}^n \to \mathbf{R}$ di classe $C^2(A)$ con $x_0 \in A^\circ$ abbiamo una formula di Taylor di secondo ordine

$$f(x_0 + h) = f(x_0) + \langle \nabla f(x_0), h \rangle + \frac{1}{2}\langle h, H_f(x_0)h \rangle + o(||h||^2) \quad \text{per} \quad h \to 0 \in \mathbf{R}^n \qquad (3.24)$$

dove $H_f(x_0)$ è la *matrice hessiana di f in x_0*; cioè

$$H_f(x_0) = \begin{bmatrix} D^2_{x_1 x_1} f & \cdots & D^2_{x_n x_1} f \\ \vdots & \ddots & \vdots \\ D^2_{x_1 x_n} f & \cdots & D^2_{x_n x_n} f \end{bmatrix}_{x = x_0} = \begin{bmatrix} \longleftarrow & \nabla(D_{x_1} f(x_0)) & \longrightarrow \\ & \vdots & \\ \longleftarrow & \nabla(D_{x_n} f(x_0)) & \longrightarrow \end{bmatrix} \qquad (3.25)$$

dove la hessiana è simmetrica per il *Teorema di Schwarz* essendo $f \in C^2(A)$.

Osservazione 13: *Possiamo riscrivere la formula di Taylor di secondo ordine come*

$$f(x_0 + h) = f(x_0) + Lh + \frac{1}{2}Q(h) + o(||h||^2) \quad \text{per} \quad h \to 0 \in \mathbf{R}^n \qquad (3.26)$$

dove $L = df(x_0)$ è lineare e $Q(h) = \frac{1}{2}\langle h, H_f(x_0)h \rangle$ è una forma quadratica. Quindi localmente lo spostamento dal grafico di f dal suo piano tangente basata in $(x_0, f(x_0))$ è controllata da $Q(h)$.

Questa osservazione è alle base delle questioni di estremi locali e convessità per funzioni scalari di più variabili. In particolare, essendo che $H_f(x_0)$ è simmetrica a valori reali sappiamo che tutti gli *autovalori* λ_j sono reali.

Osservazione 14: (Caratterizzazione della hessiana) *Sia $f \in C^2(A)$ con $x_0 \in A^\circ$. Ordiniamo gli autovalori di $H_f(x_0)$ come $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$. Si ha:*

$$H_f(x_0) \text{ sia} \begin{cases} \text{positiva definita } \Leftrightarrow \lambda_1 > 0 \\ \text{negativa definita } \Leftrightarrow \lambda_n < 0 \\ \text{positiva semidefinita } \Leftrightarrow \lambda_1 = 0 \\ \text{negativa semidefinita } \Leftrightarrow \lambda_n = 0 \\ \text{indefinita } \Leftrightarrow \exists\ \lambda_j < 0, \lambda_k > 0 \end{cases} \tag{3.27}$$

Notiamo infine che anche nel caso generale di $f : A \subseteq \mathbf{R}^n \to \mathbf{R}^m$ di classe $C^k(A; \mathbf{R}^m)$ con $k \in \mathbf{N}$ c'è una formula di Taylor di ordine k. Scrivendola componente per componente dove $f = (f_1, \ldots, f_m)$ abbiamo

$$f_i(x_0 + h) = \sum_{|\alpha| \leq k} \frac{1}{\alpha!} D_x^\alpha f_i(x_0) h^\alpha + o(||h||) \text{ per } h \to 0 \in \mathbf{R}^n, \quad i = 1, \ldots, n. \tag{3.28}$$

dove la derivata D_x^α è definita in (3.6) e si definisce per un multi-indice $\alpha = (\alpha_1, \ldots, \alpha_n)$

$$\alpha! = \alpha_1! \alpha_2! \cdots \alpha_n! \text{ e } h^\alpha = h_1^{\alpha_1} h_2^{\alpha_2} \cdots h_n^{\alpha_n}. \tag{3.29}$$

Come nel caso di funzioni da $\mathbf{R}$ in $\mathbf{R}$ ci sono formule con resto di Lagrange e resto in forma integrale (vedi il libro di De Marco per casi molto generali).

3.4 Estremi liberi

Ricordiamo che una funzione $f : A \subseteq \mathbf{R}^n \to \mathbf{R}$ (a valori **scalari**) è detto di avere *massimo locale in* $x_0 \in A$ se e solo se esite un intorno U di x_0 tale che

$$f(x_0) \geq f(x) \quad \forall x \in U \cap A \tag{3.30}$$

dove si gira il verso $f(x_0) \leq f(x)$ per un *minimo locale*. Abbiamo una generalizzazione del Teorema di Fermat.

Osservazione 15: (Condizioni necessarie) *Sia $f : A \subset\rightarrow \mathbf{R}$ con A aperto. Se f ha estremo locale in x_0 allora f differenziabile in x_0 implica $\nabla f(x_0) = 0$; cioè x_0 è un punto stazionario (critico) per f.*

Dato il fatto che il piano tangente al grafico di f in un estremo locale è orizzonatale è chiaro che la natura di un punto critico per una funzione due volte differenziabile dipende sulla hessiana.

Osservazione 16: (Condizioni sufficienti) *Sia $f : A \subset\rightarrow \mathbf{R}$ con A aperto. Se f ha un punto critico in x_0 e f è due volte differenziabile in x_0 allora*

$$H_f(x_0) \begin{cases} \text{positiva definita} \\ \text{negativa definita} \\ \text{indefinita} \end{cases} \Longrightarrow x_0 \text{ e}' \text{ un punto di} \begin{cases} \text{minimo locale} \\ \text{massimo locale} \\ \text{sella} \end{cases} \quad \text{per} \quad f$$

Ricordiamo che un punto critico si chiama *punto di sella* quando per ogni intorno U di x_0 esistano punti $x_-, x_+ \in U$ tali che $f(x_-) < f(x_0) < f(x_+)$. Nel caso in cui la hessiana è semidefinita serve un ulteriore analisi (vedi paragrafo 5.8 di [1]).

Esercizio 11: Sia A chiuso e limitato. Per ogni $x_0 \in A$ mostrare che $d_{x_0}(x) := d(x, x_0) = ||x - x_0||$ ammette massimo in qualche punto $x_M \in \partial A$. Mostrare inoltre che il diametro di A è realizzato per punti al bordo; cioè esistono $x_1, x_2 \in \partial A$ t.c. $\text{diam}(A) = ||x_1 - x_2||$.

References

[1] M. Salvatori e M. Vignati - *Appunti di Analisi Matematica II*, Dipartimento di Matematica, Università di Milano, disponibile in rete alla pagina `http://www.mat.unimi.it/users/mauras/appunti.html`.