Analisi II

Paolo Bettelini

Contents

1 Spazi metrici

1.1 Definizioni

L'insieme vuoto e l'insieme X sono aperti e chiusi.

Proposition
L'unione di aperti non numerabile è aperta, mentre l'intersezione è aperta solo se finita.

Proof
Per dimostrare quest'ultima lo facciamo su due insiemi e il resto è per induzione. Prendiamo un punto nell'interezione e prendiamo le due bolle dentro gli insiemi centrate nel punto. Siccome hanno lo stesso centro la loro intersezione è sempre una bolla di raggio il minore fra i due.

La metrice discreta può generare una bolla che è un singoletto.

Proposition
L'unione di chiusi finiti è chiusa. L'intersezione qualsiasi è chiusa.

Ogni singoletto è chiuso. Per dimostrarlo mostriamo che nel complementare esiste una bolla che non interseca il punto (vero per proprietà di Hausdorff).

Tutti i punti di accumulazione sono dei punti aderenti. Tutti i punti di un sottoinsieme sono aderenti per il sottoinsieme. Ogni punto o è di accumulazione o è isolato.

Se x_0 è aderente ad E, x_0 può essere un punto di E oppure no. Se x_0 è punto di accumulazione per E, in ogni bolla centrata in x_0 cadono inifniti punti.

Proposition

E° è aperto. E è aperto se e solo se $E = E^\circ$. E° è il più grande aperto contenuto in E.
$\overline{E}$ è chiuso. E è chiuso se e solo se $E = \overline{E}$. La chiusura di E è il più piccolo chiuso contenente E.

Proof per l'interno

Dimostriamo che E° è aperto. Sia $x_0 \in E^\circ$. un punto interno ad E, quindi esiste una bolla centrata in tale punto che è contenuta in E. Prendiamo un altro punto y in questa bolla. Possiamo costruire una inner bolla centrata in y con un raggio sufficientemente piccolo da rimanere nella bolla più grande. Quindi il punto y è a sua volta interno, quindi tutta la bolla centrata in x_0 è in E° e quindi è aperto.
Dimostriamo ora che se E è aperto allora $E = E^\circ$ (l'altra implicazione è ovvia). Per fare ciò dimostriamo che E° è il più grande aperto in E. Osserviamo che E° fa parte della famiglia degli aperti di X contenuti in E. Sia A un aperto contenuto in E. VOglio dimostrare che $A \subseteq E^\circ$. Sia $x_0 \in A$. A èunione di bolle quindi esiste unr aggio tale che la bolla centrata in x_0 di tale raggio è contenuta in A che è contenuto in E. Quindi, x_0 è interno ad E e $x_0 \in E^\circ$ e $A \subseteq E^\circ$. Supponiamo ora che E sia aperto. Allora E fa parte della famiglia degli aperti di X contenuti in E. Devo avere $E \subseteq E^\circ$. Dato che $E^\circ \subseteq E$ allora $E^\circ = E$.

Dire che un insieme è dentro in un altro significa dire che la sua chiusura coincide con l'insieme. Tipo la chiusura di Q è R quindi Q è denso in R.

Definizione Limitato

Se è contenuto in una bolla

Definizione Diametro

è il sup della metrica su tutte le coppie.

Definizione Ricoprimento

Sia E un sottoinsieme di uno spazio metrico X. Una famiglia

$$\{G_\alpha\}_{\alpha \in A}$$

è un ricoprimento apert di E se

$$E \subseteq \bigcup_{\alpha \in A} G_\alpha$$

Definizione Sottoricoprimento

Un Sottoricoprimento di

$$\{G_\alpha\}_{\alpha \in A}$$

è una sottofamiglia di G_α tale che continua a ricoprire. Cioè ne scarto alcuni ma deve comunque rimanere una copertura.

Definizione Compatto

Uno spazio metrico X è compatto se ogni ricoprimento aperto di E ammette un sottoricoprimento finito.

Ogni insieme finito è compatto.

Teorema

Sia X uno spazio metrico e E un sottoinsieme di X compatto.

1. E è limitato;
2. E è chiuso;
3. Ogni sottoinsieme infinito di E ha almeno un punto di accumulazione in E.

Proof

1. Consideriamo $\{B_1(x) \mid x \in E\}$ che è un ricoprimento aperto di E. Siccome E è compatto esiste un sottoricoprimento finito aperto di E, ossia $x_1, \ldots, x_n \in E$ tali che
$$E \subseteq \bigcup_{i=1}^{n} B_1(x_i)$$
Posto
$$R = 1 + \max_{i=1,\ldots,n} d(x_i, x_1)$$
Allora la bolla di raggio R centrata in x_1 contiene E, quindi E è limitato.
2. Supponiamo che non sia chiuso. Allora esiste $y \in E'$ ma $y \notin E$. Vogliamo costruire un ricoprimento aperto di E che non ammette sottoricoprimento finito. Sia $r(x) = \frac{1}{2}d(x, y)$ per ogni $x \in X$. Se $x \in E$ allora $r(x) > 0$ perchè $y \notin E$. Abbiamo il ricoprimento
$$\{B_{r(x)}(x) \mid x \in E\}$$
Ma per la compattezza esisterebbe un sottoricoprimento finito, cioè $x_1, \ldots, x_n \in E$ tali che
$$E \subseteq \bigcup_{i=1}^{n} B_{r(x_i)}(x_i)$$
Sia ora $R = \min_{i=1,\ldots,n} r(x_i)$. Allora $R > 0$ e la bolla $B_R(y)$ non interseca nessuna delle $B_{r(x_i)}(x_i)$, assurdo poiché y è punto di accumulazione.
3. Sia F un sottoinsieme infinito di E. Supponiamo che F non abbia punti di accumulazione in E. Allora ogni punto di E ha una bolla che interseca F in al più un punto. Queste formano un ricoprimento aperto di E. Ma se esistesse un sottoricoprimento finito, F sarebbe finito, assurdo.

Proposition

Sia $E \subseteq X$ compatto. Se $F \subseteq E$ è chiuso allora F è compatto.

Proof

Sia $\{G_\alpha\}_{\alpha \in A}$ un ricoprimento aperto di F. Dobbiamo aggiungere degli insiemi aperti per coprire il resto. Siccome F è chiuso, $X \setminus F$ è aperto. Quindi $\{G_\alpha\}_{\alpha \in A} \cup \{X \setminus F\}$ è un ricoprimento aperto di E. Per la compattezza di E esiste un sottoricoprimento finito, che escludendo $X \setminus F$ è un sottoricoprimento finito di F.

Se $F \subseteq X$ è chiuso, ed $E \subseteq X$ è compatto, allora $F \cap E$ è compatto.

Teorema Teorema dell'intersezione finita

Sia $\{E_\alpha\}_{\alpha \in A}$ una famiglia di compatti tale che ogni intersezione finita è non vuota. Allora

$$\bigcap_{\alpha \in A} E_\alpha \neq \emptyset$$

Proof

Supponiamo che l'intersezione sia vuota. Allora e sia $E_{\overline{\alpha}}$ un compatto fissato nella famiglia.

$$\begin{aligned} & E_{\overline{\alpha}} \cap \left(\bigcap_{\alpha \neq \overline{\alpha}} E_\alpha \right) = \emptyset \\ \implies & E_{\overline{\alpha}} \subseteq \left(\bigcap_{\alpha \neq \overline{\alpha}} E_\alpha \right)^c = \bigcup_{\alpha \neq \overline{\alpha}} E_\alpha^c \end{aligned}$$

$\{E_\alpha^c\}_{\alpha \neq \overline{\alpha}}$ è un ricoprimento aperto di $E_{\overline{\alpha}}$. Esistono quindi $\alpha_1, \ldots, \alpha_n \neq \overline{\alpha}$ tali che

$$\begin{aligned} & E_{\overline{\alpha}} \subseteq \bigcup_{i=1}^n E_{\alpha_i}^c = \left(\bigcap_{i=1}^n E_{\alpha_i} \right)^c \\ \implies & E_{\overline{\alpha}} \cap \left(\bigcap_{i=1}^n E_{\alpha_i} \right) = \emptyset \end{aligned}$$

assurdo.

Corollario caso particolare

Sia $\{E_n\}_{n \in \mathbb{N}}$ una famiglia di compatti tale che

$$E_{n+1} \subseteq E_n$$

Allora

$$\bigcap_{n \in \mathbb{N}} E_n \neq \emptyset$$

Teorema Teorema di Heine-Borel

Sia $E \subseteq R^n$ con la metrica euclidea. Allora E è compatto se e solo se E è chiuso e limitato.

Lemma

Sia $\{I_k\}_{k \in \mathbb{N}}$ una famiglia di intervalli $I_k = [a_k, b_k]$ tali che $I_k \supseteq I_{k+1}$. Allora

$$\bigcap_{k \in \mathbb{N}} I_k \neq \emptyset$$

Proof

Gli intervalli sono annidati, quindi a_k è crescente e b_k è decrescente e $a_k \leq b_k$. In particolare $a_k \leq b_i$. Consideriamo l'insieme $E = \{a_k \,|\, k \in \mathbb{N}\}$. E è limitato superiormente, e ammette supremum x. Per definizione $x \geq a_k$. Ma $a_k \leq b_i$ per tutte le i. Quindi, $x \leq b_i$ per ogni i. Allora

$$x \in I_n \implies x \in \bigcap I_k$$

Definizione

Siano $a, b \in \mathbb{R}^n$ con $a_i < b_i$ per ogni $i = 1, \ldots, n$. Un rettangolo chiuso è il prodotto cartesiano

$$[a_1, b_1] \times [a_2, b_2] \times \ldots \times [a_n, b_n]$$

che indichiamo con $[a, b]$.

Lemma

Sia $\{R_k\}_{k \in \mathbb{N}}$ una famiglia di rettangoli chiusi tali che $R_k \supseteq R_{k+1}$ per ogni k. Allora

$$\bigcap_{k \in \mathbb{N}} R_k \neq \emptyset$$

Proof

Siccome

$$R_k = I_{k,1} \times I_{k,2} \times \ldots \times I_{k,n}$$

possiamo applicare il primo lemma e quindi

$$\exists y_i \in \bigcap_{k \in \mathbb{N}} I_{k,i}$$

Il punto $y = (y_1, \ldots, y_n)$ è in ogni R_k.

Lemma Lemma 3

In $\mathbb{R}^n$ con la metrica euclidea ogni rettangolo è compatto.

Proof Lemma 3

Sia $R = [a, b]$ un rettangolo e supponiamo che non sia compatto. Sia $\{G_\alpha\}_{\alpha \in A}$ un ricoprimento aperto di R che non ammette sottoricoprimento finito. Vogliamo adesso dimezzare ambo i lati (quindi n tagli). Abbiamo adesso 2^n rettangoli.

$$[a_i, b_i] = [a_i, c_i] \cup [c_i, b_i], \quad c_i = \frac{a_i + b_i}{2}$$

Il diametro di R è $||b - a||$. Il diametro di ogni rettangolo ottenuto è la metà. Almeno uno di questi rettangoli ha la proprietà di non ammettere sottoricoprimento finito. Lo chiamiamo R_1. Iterando il procedimento otteniamo una successione di rettangoli

$$R \supseteq R_1 \supseteq R_2 \supseteq \ldots$$

con diametro che tende a zero e che non ammettono sottoricoprimento finito, il diametro di R^k è dato da $\frac{1}{2^k}||b - a||$. Per il lemma precedente esiste $x \in \bigcap_{k \in \mathbb{N}} R_k$. Siccome $R_k \subseteq R$ per ogni k, $x \in R$. Siccome $\{G_\alpha\}_{\alpha \in A}$ è un ricoprimento di R, esiste $\alpha_0 \in A$ tale che $x \in G_{\alpha_0}$. G_{α_0} è aperto, quindi esiste $r > 0$ tale che $B_r(x) \subseteq G_{\alpha_0}$. Scegliamo k sufficientemente grande tale che $2^{-k}||b - a|| < r$. Ma il diametro di R_k è minore di r, quindi $R_k \subseteq B_r(x)$. Quindi $R_k \subseteq G_{\alpha_0}$, assurdo perchè R_k non ammette sottoricoprimento finito.

Proof Heine-Borel

Dobbiamo dimostrare solo che se E è chiuso e limitato allora è compatto. Siccome E è limitato esiste M tale che $||x|| < M$ per ogni $x \in E$. Quindi,

$$E \subseteq [-M, M] \times [-M, M] \times \ldots \times [-M, M] = R$$

E è un chiuso contenuto in un compatto, quindi è compatto.

Teorema Teorema di Bolzano-Weierstrass
Ogni sottoinsieme infinito e limitato di $\mathbb{R}^n$ ha almeno un punto di accumulazione.

Proof Teorema di Bolzano-Weierstrass

Definizione Insiemi separati
Sia (X, d) uno spazio metrico e $A, B \subseteq X$ due sottoinsiemi. Diciamo che A e B sono separati se

$$A \cap \overline{B} = \emptyset \wedge \overline{A} \cap B = \emptyset$$

Devono sicuramente essere disgiunti, ma non basta. Serve che nessun punto di uno dei due insiemi è punto di accumulazione dell'altro.

Definizione
Sia (X, d) uno spazio metrico e $E \subseteq X$. E è connesso se non può essere scritto come unione di due sottoinsiemi non vuoti e separati.

I sottoinsiemi connessi di $\mathbb{R}$ sono tutti e soli gli intervalli.

Uno spazio metrico è connesso se e solo se l'unico sottoinsieme non vuoto di X che è anche aperto e chiuso è X stesso. (Dimostrazione per esercizio).

R^n con la metrica euclidea è connesso. (Dimostrazione per esercizio non proprio banale).

1.2 Successioni in spazi metrici

Mettere la definizione di convergenza ma con $d(x_m, y) < \varepsilon$. Oppure $x_m \in B_\varepsilon(y)$.

In particolare la successione metrica converge se e solo se $d(x_m, y) \to 0$ secondo la convergenza reale.

Il limite è unico per proprietà di Hausdorff.

Proposition
Sia (X, d) uno spazio metrico e $E \subseteq X$ e sia y un punto di accumulazione per E. Allora esiste una successione $\{x_n\} \subseteq E \setminus \{y\}$ che converge ad y. In particolare, E è chiuso se e solo se per ogni successione $\{x_n\} \subseteq E$ che converge ad y allora $y \in E$.

Proof
Dato che $y \in E'$, $\forall x_m \in \mathbb{N}$, esiste x_m tale che $x_m \in B_{\frac{1}{m}}(y) \cap E$ e $x_m \neq y$. La successione così costruita converge ad y. Infatti, $d(x_m, y) < \frac{1}{m} \to 0$.

Proposition
Sia (X, d) uno spazio metrico e sia $\{x_n\}$ una successione convergente in X. Una condizione necessaria per la convergenza è che ogni sottosuccesione converga allo stesso limite. La condizione

sufficiente è che ogni sottosuccessione ammetta una sottosuccessione che converge allo stesso limite.

Definizione Compattezza sequenziale

Uno spazio metrico X è sequenzialmente compatto se ogni successione in X a valori in E ammette una sottosuccessione convergente ad un punto di E.

Proposition Equivalenza compattezza

E is compact is and only if E is sequentially compact.

Questa c'è solo negli spazi metrici.

Proof

($\Longrightarrow$) Sia $\{x_n\}$ una successione in E. Consideriamo $F = \{x_n \,|\, n \in \mathbb{N}\}$. Se F è finito, esiste un elemento che compare infiniti volte e la successione costante converge a tale elemento. Se F è infinito, per la compattezza F ammette un punto di accumulazione, $y \in E$. Costruiamo una sottosuccessione che converga ad y. Scegliamo x_m tale che $d(x_m, y) < 1$. Scegliamo x_{m_2} tale che $d(x_{m_2}, y) < \frac{1}{2}$ e $m_2 > m_1$, e così via. La sottosuccessione così costruita converge ad y in quanto $d(x_{m_k}, y) < \frac{1}{k} \to 0$.

($\Longleftarrow$) XXX

Ogni successione convergente è di Cauchy.

Per esempio con la metrica discreta una successione è convergente se e solo se è definitamente costante, che è equivalente ad essere di Cauchy, quindi è completo.

Nel caso dei razionali nei reali con metrica euclidea, consideriamo la radice di due che è un punto di accumulazione per i razionali. Esiste una successione di razionali che converge a radice di due, quindi è di Cauchy. Ma essa non può convergere in Q, altrimenti convergerebbe anche in R e avrebbe due limiti. Tuttavia è una successione di Cauchy in Q perché è convergente in R e quindi è di Cauchy in R. (La condizione è la medesima). Quindi Q non è completo.

Definizione Spazio completo

Uno spazio metrico (X, d) è completo se ogni successione di Cauchy in X converge ad un punto di X.

Teorema

R^n con la metrica euclidea è completo.

Proof

Sia $\{x_n\}$ una successione di Cauchy in R^n. Scriviamo $E_n = \{x_k \,|\, k \geq n\}$. Notiamo che $E_n \supseteq E_{n+1}$. Ponendo la chiusura $\overline{E_n} \supseteq \overline{E_{n+1}}$. Inoltre, E_n è limitato e $\text{diam} E_n \to 0$. Infatti, dato $\varepsilon > 0$ esiste N tale che per ogni $m, n \geq N$ $d(x_n, x_m) < \varepsilon$. Notiamo inoltre che

$$\text{diam} E_n = \sup\{d(x_m, x_k)\} < \varepsilon$$

Dimostrazione per esercizio vale che $\text{diam} F = \text{diam} \overline{F}$. Quindi, $\text{diam} \overline{E_n} \to 0$. Adesso $\{\overline{E_n}\}$ è una successione di compatti in quanto chiusi e limitati, annidati. Quindi

$$E \triangleq \bigcap_{n \in \mathbb{N}} \overline{E_n} \neq \emptyset$$

Siccome diam$E = 0$ o è vuoto o contiene un solo punto, quindi contiene un solo punto $E = \{y\}$. Mostriamo che $x_n \to y$. Abbiamo $d(x_n, y) \leq \text{diam}\overline{E_n} \to 0$.

Teorema

Sia (X, d) uno spazio metrico compatto. Allora X è completo.

Proof

Sia $\{x_n\}$ una successione di Cauchy in X. Siccome è compatto è compatto per successioni, quindi esiste una sottosuccessione $\{x_{n_k}\}$ che converge ad un punto $y \in X$. Mostriamo che $x_n \to y$. Dato $\varepsilon > 0$ esiste N_0 tale che per ogni $m, n \geq N_0$ $d(x_n, x_m) < \frac{1}{2}\varepsilon$. Per la convergenza di $\{x_{n_k}\}$ esiste K tale che per ogni $k \geq K$ $d(x_{n_k}, y) < \frac{1}{2}\varepsilon$. Scegliamo $\overline{N} = \max\{N_0, n_K\}$. Allora per ogni $n \geq \overline{N}$ si ha

$$d(x_n, y) \leq d(x_n, x_{n_K}) + d(x_{n_K}, y) < \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon$$

Sia X uno spazio metrico completo, $Y \subseteq X$. Y è completo se e solo se Y è chiuso in X.

Teorema

E sequenzialmente comaptto implica E compatto.

Proof

Sia $\{G_\alpha\}_{\alpha \in A}$ un ricoprimento aperto di E. Esiste $\delta > 0$ tale che $\forall x \in E$ esiste $\overline{\alpha}$ tale che $B_\delta(x) \subseteq G_{\overline{\alpha}}$.

1. *claim 1:* $\forall m \in \mathbb{N}$, esiste x_m tale che $B_{1/m}(x_m)$ non è sottoinsieme di G_α per tutte le α. $\{x_m\}$ è una successione in E e quindi posso estrarre una sottosuccessione convrgente $x_{m_k} \to p \in E$. Esiste $\hat{\alpha}$ tale che $p \in G_{\hat{\alpha}}$. $G_{\hat{\alpha}}$ è aperto e quindi esiste una $\varepsilon > 0$ tale che $B_\varepsilon(p) \in G_{\hat{\alpha}}$. Ma $x_{m_k} \to p$ quindi con k sufficientemente grande

   $$B_{1/m_k}(x_{m_k}) \subseteq B_\varepsilon(p) \subseteq G_{\hat{\alpha}}$$

   che è assurdo lightning.
2. *claim 2:* E è contenuto nel'unione di un numero finito di bolle di raggio δ centrate in punto di E. Per assurdo, sia $x_1 \in E$. Sicuramente $B_\delta(x_1)$ non ricopre E quindi esiste $x_2 \in E \backslash B_\delta(x_1)$. Ma assieme $B_\delta(x_1) \cup B_\delta(x_2)$ non ricoprono E, quindi esiste un $x_3 \in E \backslash (B_\delta(x_1) \cup B_\delta(x_2))$ e così via. La successione $\{x_m\}$ deve ammettere una sottosuccessione convergente. Ma $d(x_i, x_j) \geq \delta$ se $i \neq j$ quindi la sucessione $\{x_m\}$ non è di Cauchy Lightning. Quindi $E \subseteq B_\delta(x_1) \cup B_\delta(x_2) \cup \cdots$.

In realtà abbiamo mostrato anche la terza.

Teorema

Sia X uno spazio metrico. Sono equivalenti:

1. X è compatto;
2. X è sequenzialmente compatto;
3. *limit point compact:* ogni sottoinsieme infinito di X ha almeno un punto di accumulazione.

Solo negli spazi metrici.

1.3 Funzioni

Definizione

Siano $(X_1, d_1), (X_2, d_2)$ due spazi metrici e sia $E \subseteq X_1$. Sia $f\colon E \to X_2$ e $p \in E'$. Diciamo che $l \in X_2$ è limite di $f(x)$ per $x \to p$ e diciamo

$$\forall \varepsilon > 0, \exists \delta > 0 \,|\, x \in E \wedge 0 < d_1(x_1, p) < \delta \implies d_2(f(x), l) < \varepsilon$$

Equivalentemente $\forall \varepsilon > 0$ esiste $\delta > 0$ tale che

$$f((B_\delta(p) \cap E) \backslash \{p\}) \subseteq B_\varepsilon(l)$$

Proposition

Sia $f\colon E \subseteq X_1 \to X_2$. Allora $f(x) \to l$ per $x \to p$ se e solo se $f(x_n) = l$ per ogni successione $\{x_n\}$ tale che $x_n \in E$e $x_n \neq p$ per tutte le n e $x_n \to p$.

Valgono i medesimi teoremi tipo l'unicità del limite e i teoremi di permanenza del segno, confronto etc.

Proposition

Sia $f\colon E \subseteq X \to \mathbb{R}^n$ per $n > 1$. Allora

$$f(x) \to l \iff f_i(x) \to l_i$$

per $x \to p$.

Proof Sketch

Conderiamo la norma per tutte le i

$$|f_i(x) - l_i| \leq ||f(x) - l|| \leq \sum_k |f_k(x) - l_k|$$

Definizione Continuità

Siano $(X_1, d_1), (X_2, d_2)$ due spazi metrici, $f\colon E \subseteq X_1 \to X_2$, $p \in E$. Diciamo che f è *continua* in p se $\forall \varepsilon > 0$ esiste $\delta > 0$ tale che

$$\forall x \in E \cap B_\delta(p) \implies f(x) \in B_\varepsilon((f(p)))$$

Euivalentemente $\forall \varepsilon > 0$ esiste $\delta > 0$ tale che $x \in E$ e $d_1(x, p) < \delta$ implica che $d_2(f(x), f(p)) < \varepsilon$. Oppure ancora $(f(B_\delta(p) \cap E)) \subseteq B_\varepsilon(f(p))$.

Se p è un punto isolato di E allora $\exists r > 0$ tale che $B_r(p) \cap E = \{p\}$. Scegliendo $\delta \leq r$ la definizione di continuità è automaticamente soddisfatta. Se p non è i solato, allora è un punto di accumulazione per E. In questo caso f è continua in p e vale che $f(x) \to f(p)$ per $x \to p$.

Proposition

f è continua in p se e solo se

$$\lim_{x \to p} f(x_n) = f(p)$$

per ogni successione $\{x_n\}$ tale che $x_n \in E$ per tutte le n e $x_n \to p$.

Definizione

Sia $f\colon E \subseteq X_1 \to X_2$. Diciamo che f è continua nell'insieme E se f è continua in ogni punto di E.

Proposition

Siano $(X_1, d_1), (X_2, d_2)$ spazi metrici e sia $f\colon X_1 \to X_2$. Allora f è continua in X se e solo se $f^{-1}(V)$ è aperto in X_1 per tutti i V aperti in X_2.

Proof

($\Longrightarrow$) Sia V un aperto di X_2. Se $f^{-1}(V) = \emptyset$ in questo caso abbiamo finito. Altrimenti, sia $p \in f^{-1}(V)$, cioè $f(p) \in V$. Essendo V aperto, riesco a trovare

$$B_\varepsilon(f(p)) \in V$$

Ma f è continua quindi riesco anche a trovare $\delta > 0$ tale che

$$f(B_\delta(p)) \subseteq B_\varepsilon(f(p))$$

Quindi $B_\delta(p) \subseteq f^{-1}(V)$ quindi p è un punto interno a $f^{-1}(V)$. Per l'arbitrarietà di p segue che $f^{-1}(V)$ è aperto.

($\Longleftarrow$) Sia $p \in X$ e dimostriamo che f è continua in p. Sia $\varepsilon > 0$ fissato. $B_\varepsilon(f(p))$ è un aperto di X_2. $f^{-1}(B_\varepsilon(f(p)))$ è un aperto di X_1 e $p \in f^{-1}(B_\varepsilon(f(p)))$ e quindi esiste $\delta > 0$ tale che

$$B_\delta(p) \subseteq f^{-1}(B_\varepsilon(f(p)))$$

cioè

$$f(B_\varepsilon(p)) \subset B_\varepsilon(f(p))$$

che è la definizione di continuità.

Siccome $f^{-1}(E^c) = \left(f^{-1}(E)\right)^c$ allora f è continua se e solo se $f^{-1}(C)$ è chiuso in X_1 per ogni chiuso in $C \in X_2$. Molto utile.

In generale le funzioni continue non mandano aperti in aperti. Per esempio $f\colon \mathbb{R} \to \mathbb{R}$ data da $x \to x^2$. Abbiamo che

$$f((-1,1)) = [0,1)$$

Definizione Funzione aperta

Una funzione viene detta *aperta* se $f(U)$ è aperto in X_2 per tutti gli insiemi U aperto om X_1. Analogamente funzione chiusa.

Sia $f\colon (X,d) \to \mathbb{R}^n$ con $n > 1$ è continua se e solo se tutte le sue componenti sono continue.

Proposition

Siano $(X_1, d_1), (X_2, d_2)$ spazi metrici, $f\colon X_1 \to X_2$ una funzione continua. Se X_1 è compatto, allora $f(X_1)$ è compatto.

Proof

Sia $\{G_\alpha\}_{\alpha \in A}$ un ricoprimento aperto di $f(X_1)$. Consideriamo $\{f^{-1}(G_\alpha)\}_{\alpha \in A}$ che sono degli aperti. Queste preimmagini sono un ricoprimento di X_1, che è compatto e quindi posso estrarre un sottoricoprimento finito $f^{-1}(G_{\alpha_1}), \cdots, f^{-1}(G_{\alpha_n})$. Vogliamo mostrare che $\{G_{\alpha_1}, \cdots, G_{\alpha_n}\}$ sono un ricoprimento di $f(X_1)$.

$$f(X_1) = f\left(\bigcup_{i=1}^{n} f^{-1}(G_{\alpha_i})\right) = \bigcup_{i=1}^{n} f\left(f^{-1}(G_{\alpha_i})\right) \subseteq \bigcup_{i=1}^{n} G_{\alpha_i}$$

Teorema Teorema di Weierstrass

Sia (X, d) uno spazio metrico compatto e sia $f\colon X \to \mathbb{R}$ una funzione continua. Allora, $\exists x_1, x_2 \in X$ tali che

$$f(x_1) \leq f(x) \leq f(x_2), \quad \forall x \in X$$

cioè f possiede massimo e minimo assoluto.

Proof Teorema di Weierstrass

$f(X)$ è compatto in $\mathbb{R}$, quindi è chiuso e limitato. Siccome limitato, $f(X)$ ammette infimum e supremum reali. Siccome $\inf f(x)$ e $\sup f(x)$ appartengono a $\overline{f(x)}$ e $f(x)$ è chiuso, appartengono allora ad $f(X)$ e quindi sono massimi e minimi.

Teorema Teorema da compatto ad Hausdorff

Siano $(X_1, d_1), (X_2, d_2)$ spazi metrici con X_1 compatto e $f\colon X_1 \to X_2$ continua. Allora, f è chiusa.

In realtà questo funziona con domini compatti e codomini di Hausdorff.

Proof Teorema da compatto ad Hausdorff

Sia C un chiuso di X_1. Voglio dimostrare che $f(C)$ è un chiuso di X_2. Sappiamo che C è chiuso in un compatto, quindi è compatto. La funzione è continua e quindi $f(C)$ è compatto. Siccome i compatti sono chiusi allora è chiuso.

Corollario

Sia $f\colon (X_1, d_1) \to (X_2, d_2)$ continua, X_1 compatto e f biunivoca. Allora, f^{-1} è continua.

Proof

Dobbiamo mostrare che $(f^{-1})^{-1}(C)$ è chiuso per ogni C chiuso di X_2. Ma questo coincide con $f(C)$ che è chiusa per il teroema da compatto ad Hausdorff.

Teorema

Sia $f\colon (X_1, d_1) \to (X_2, d_2)$ continua e sia $E \subseteq X$ connesso. Allora $f(E)$ è connesso.

Proof

Supponiamo che $f(E)$ non sia connesso. Esistono quindi due sottoinsiemi non vuoti disgiunti e separati tali che

$$f(E) = A \cup B$$

Poniamo $F = f^{-1}(A) \cap E$ e $G = f^{-1}(B) \cap E$. Sicuramente $F, G \neq \emptyset$ e $E = F \cup G$. Vogliamo mostrare che E e G sono separati. Siccome $A \subseteq \overline{A}$ vale anche $f^{-1}(A) \subseteq f^{-1}(\overline{\overline{A}})$. L'applicazione f è continua e la chiusura di A è un chiuso. Quindi la preimmagine del chiuso $\overline{A}$ è un chiuso. Consideriamo ora

$$\overline{F} \subseteq \overline{f^{-1}(A)} = f^{-1}(\overline{A})$$

perché f è continua se $\overline{A}$ è chiuso. Quindi $\overline{F} \subseteq f^{-1}(\overline{A})$ che implica $f(\overline{F}) \subseteq \overline{A}$. D'altro canto $f(G) \subseteq B$ e $\overline{A} \cap B \neq 0$, e quindi $\overline{F} \cap G \neq 0$ perché altrimenti vi sarebbe un elemento sia in $\overline{A}$ che in B. Dovrebbe essere $f(x) \in \overline{A}$ e $f(x) \in B$ lightinng. Analogamente si dimostra che $F \cap \overline{G} = \emptyset$ cioè abbiamo scritto E come unione di due sottoinsiemi non vuoti e separati. Ma E è connesso lightning.

Definizione

Siano $(X_1, d_1), (X_2, d_2)$ spazi metrici e $f\colon X_1 \to X_2$. Allora f è uniformemente continua se $\forall \varepsilon > 0$, esiste $\delta > 0$ tale che $\forall x, y \in X_1$

$$d_1(x, y) < \delta \implies d_2(f(x), f(y)) < \varepsilon$$

Teorema Theorema di Heine-Cantor

Siano $(X_1, d_1), (X_2, d_2)$ spazi metrici e $f\colon (X_1, d_1) \to (X_2, d_2)$ continua e X_1 compatto. Alorra, f è uniformemente continua.

La dimostrazione è la medesima rispetto al caso banale.

Definizione Funzione di Lipschitz

Siano $(X_1, d_1), (X_2, d_2)$ spazi metrici, $f\colon X_1 \to X_2$. Diciamo che f è *lipschitz-continua* o *lipschitziana* se $\exists \alpha > 0$ tale che

$$d_2(f(x), f(y)) \leq \alpha d_1(x, y)$$

per tutte le $x, y \in X_1$.

Proposition

Se f è Lipschitz-continua, allora è uniformemente continua.

Definizione

Siano $(X_1, d_1), (X_2, d_2)$ spazi metrici e $f\colon (X_1, d_1) \to (X_2, d_2)$. Diciamo che f è una *contrazione* se $f \in \mathrm{Lip}_\alpha(X_1, X_2)$ con $\alpha < 1$.

Se il supremum è finito, allora questa è la miglior costante di Lipschitz (in generale)

$$\sup_{x \neq y} \frac{|f(x) - f(y)|}{|x - y|}$$

Esempio

Consideriamo $f\colon \mathbb{R} \to \mathbb{R}$ data da $f(x) = x^2$. Non è di lipschitz inquanto non è uniformemente continua. Per mostrarlo possiamo dire

$$|f(x) - f(y)| = |x^2 - y^2| = |x + y| \cdot |x - y|$$

Se restringessimo il dominio di questa funzione ad un intervallo limitato, allora sarebbe di Lipschitz, per il supremum.

Proposition

Sia $f\colon I \subseteq \mathbb{R} \to \mathbb{R}$ differenziabile. Allora, $f \in \mathrm{Lip}_\alpha(I, \mathbb{R})$ se e solo se $|f'(x)| \leq \alpha$ per tutte le x.

Proof

($\Longrightarrow$) Cominciamo con

$$
\begin{aligned}
|f'(x)| &= \left|\lim_{t\to 0} \frac{f(x+t)-f(x)}{t}\right| \\
&= \lim_{t\to 0} \left|\frac{f(x+t)-f(x)}{t}\right| \\
&= \lim_{t\to 0} \frac{|f(x+t)-f(x)|}{|t|} \\
&\leq \lim_{t\to 0} \frac{\alpha|x+t-x|}{|t|} = \alpha
\end{aligned}
$$

Possiamo togliere il limite dal modulo in quanto il modulo è una funzione continua.

($\Longleftarrow$) Per il teorema di Lagrange esiste $\theta \in (\min\{x,y\}, \max\{x,y\})$

$$
\begin{aligned}
f(x)-f(y) &= f'(\theta)(x-y) \\
|f(x)-f(y)| &= |f'(\theta)| \cdot |x-y| \\
&\leq \alpha|x-y|
\end{aligned}
$$

Esercizio

Per quali $a \leq b$ la funzione $f(t) = 1 + t - \arctan(t)$ è una contradizione in $[a,b]$. Stabiliamo quindi se la derivata è limitata

$$
f'(t) = 1 + \frac{1}{1+t^2} = \frac{t^2}{1+t^2}
$$

notiamo quindi che $0 \leq f'(t) \leq 1$. Quindi è sicuramente lipschitziana. Notiamo allora che

$$
\sup_{\mathbb{R}} |f'(t)| = 1 = \alpha
$$

Quindi per far sì che $\alpha < 1$ dobbiamo limitare il dominio ad un intervallo limitato. Quindi $-\infty < a \leq b < \infty$. Porta l'intervallo $[a,b]$ in sè? Siccome la funzione è crescente porta intervalli a intervalli di estremi $f(a)$ e $f(b)$. Mi basta quindi imporre che $f(a) \geq a$ e $f(b) \leq b$. Abbiamo quindi

$$
\begin{cases} 1 + a - \arctan a \geq a \\ 1 + b - \arctan b \leq b \end{cases} = \begin{cases} \arctan a \leq 1 \\ \arctan b \geq 1 \end{cases}
$$

e quindi $a \leq \tan 1 \leq b$. Notiamo che $f(\tan 1) = \tan 1$ quindi è un punto fisso per il teorema delle contrazioni.

Esempio

Sia $v \in \mathbb{R}^n$ e consideriamo $f\colon \mathbb{R}^n \to \mathbb{R}$ data da $f(x) = v \cdot x$. Dobbiamo studiare $|f(x) - f(y) = |v \cdot x - v \cdot y|$ usando la bilinearità del prodotto scalare ottengo $|v \cdot (x-y)|$. Per Cacuhy-Schwarz

$$
|v \cdot (x-y)| \leq ||v|| \cdot ||x-y||
$$

che è quindi di Lipschitz.

Teorema Teorema di Banach-Cacciopoli o delle contrazioni

Sia (X,d) uno spazio metrico completo e sia $f\colon X \to X$ una contrazione. Allora $\exists_{=1}\, x \in X$ tale che $f(x) = x$.

Le ipotesi sono necessarie. Togliamo per esempio la completezza e consideriamo quindi $X = (0, +\infty)$ con la contrazione $f(x) = x/2$. Questa contrazione non ha punti fissi. Togliamo invece l'ipotesi che sia una contrazione. Richiediamo solamente che sia una contrazione debole, cioè

$$d_2(f(x), f(y)) \leq f_1(x, y)$$

Consideriamo $X = [0, +\infty)$ e prendiamo $f(t) = t + e^{-t}$. La derivata è $f'(t) = 1 - e^{-t}$ che è nulla nell'origine e poi tende ad 1 dal sotto. Chiaramente non ci sono punti fissi in quanto $f(t) = t$ è come dire $e^{-t} = 0$.

Proof

Cominciamo mostrando l'esistenza del punto fisso. Sia $x_0 \in X$ un punto fissato e consideriamo la successione $x_{n+1} = f(x_n)$.

1. Mostriamo che $\{x_n\}$ è di Cauchy, quindi siccome lo spazio è completo converge. Dobbiamo mostrare che $d(x_n, x_m)$ tende a zero quando n, m crescono. Consideriamo inizialmente

$$\begin{aligned} d(x_{n+1}, x_n) &= d(f(x_n), f(x - n - 1)) \leq \alpha d(x_n, x_{n-1}) \\ &= \alpha d(f(x_{n-1}), f(x_{n-2})) \leq \alpha^2 d(x_{n-1}, x_{n-2}) \\ &\leq \alpha^n d(x_1, x_0) \end{aligned}$$

   Calcoliamo ora la distanza generica e usiamo la disuguaglianza triangolare ripetutamente per ogni step

$$\begin{aligned} d(x_{n+k}, d_n) &\leq \sum_{i=0}^{k-1} d(x_{n+k-i}, d_{n+k-i-1}) \\ &\leq d(x_1, x_0) \sum_{i=0}^{k-1} \alpha^{n+k-i} \\ &= \alpha^n d(x_1, x_0) \sum_{i=0}^{k-1} \alpha^{n+k-i-1} \\ &= \alpha^n \frac{\alpha^k - 1}{\alpha - 1} d(x_1, x_0) \\ &= \frac{\alpha^n}{1 - \alpha} d(x_1, x_0) \to 0 \end{aligned}$$

   Per sbarazzarci di k (siccome vogliamo k arbitrario e il ε nella definizione di Cauchy deve essere uniforme rispetto ad esso) maggioriamo la somma parziale della serie geometrica con il valore della serie geometrica. Siccome $0 < \alpha < 1$ il termine non esplode e la serie geometrica converge.
2. Detto x il limite di $\{x_n\}$ mostriamo che è un punto fisso di f. Consideriamo il limite per $n \to \infty$

$$\lim_{n\to\infty} f(x_n) = f\left(\lim_{n\to\infty} x_n\right) = f(x)$$

   perché f è continua.

Mostriamo ora l'unicità del punto. Supponiamo che x, y siano due punti fissi. Vogliamo mostrare che $d(x, y) = 0$. Abbiamo

$$d(x, y) = d(f(x), f(y)) \leq \alpha d(x, y)(1 - \alpha)d(x, y) \qquad \leq 0$$

siccome $1 - \alpha > 0$ ciò succede solo se $d(x, y) = 0$.

Abbiamo notato che

$$d(x_{n+k}, x_n) \leq \frac{\alpha^n}{1-\alpha} d(x_1, x_0)$$

Con $k \to \infty$ otteniamo

$$d(x, x_n) \leq \frac{\alpha^n}{1-\alpha} d(x_1, x_0)$$

quindi tende al punto fisso in maniera esponenziale.

Denotiamo $f^n = f \circ f \cdots f$. Se f è una contrazione, una qualsiasi sua iterazione è anch'essa una contrazione.

$$d(f(f(x)), f(f(y))) \leq \alpha d(f(x), f(y)) \leq \alpha^2 d(x, y)$$

Per induzione segue il resto. In generale la costante è α^n. Ci chiediamo se nel caso in cui f non sia una contrazione, una sua iterata lo possa essere.

Esempio

Per esempio $f(x) = \cos x$, che non è una contrazione in quanto il supremum della derivata è 1. Invece, $\cos(\cos(x))$ ha derivata

$$\sin(\cos(x)) \cdot \sin$$

Il suo modulo è dato da

$$|\sin(\cos(x))| \cdot |\sin x| \leq \sin(1) < 1$$

Il secondo termine può solamente essere maggiorato da 1, mentre il secondo, siccome $-1 \leq \cos(x) \leq 1$, può essere maggiorato da $\sin 1$. Quindi è una contrazione.
Con questo possiamo per esempio mostrare che il coseno ha un punto fisso, siccome una sua iterata è una contrazione.

Corollario Indebolimento del teorema delle contrazioni: teorema delle iterate contrazioni

Sia (X, d) uno spazio metrico completo e sia $f\colon X \to X$ un'applicazione tale che $\exists n \in \mathbb{N}$ tale che f^n sia una contrazione. Allora $\exists_{=1}\, x \in X$ tale che $f(x) = x$.

Proof

Mostriamo che i punti fissi di f (che sono uno solo) sono i punti fissi di f^n. Sia x un punto fisso di f. Allora $f^n(x) = f(f(\cdots(x)) = f(x) = x$. Quindi tutti i punti fissi di f sono anche punti fissi di f^n. Sia ora x tale che $f^n(x) = x$. Componendo otteniamo

$$f(f^n(x)) = f(x)$$
$$f^n(f(x)) = f(x)$$

quindi $f(x)$ è un punto fisso di f^n, ma siccome f è una contrazione ha solo un punto fisso, quindi coincidono $f(x) = x$. Quindi tutti i punti fissi di f^n sono anche punti fissi di f.

Parametrizziamo ora la funzione

Consideriamo $T\colon X \times Y \to X$ come operatore parametrizzato dai valori di Y. Fissato y imponiamo che $T(-, y) : X \to X$ sia una contrazione. Per tutte le y esiste un solo $x \in X$ tale che $T(x, y) = x$. Data la dipendenza funzionale $x = \varphi(y)$ vogliamo capire come il punto fisso dipende dal parametro. In particolare, vogliamo mostrare che φ è continua sotto alcune ipotesi.

Teorema di dipendenza del punto tfisso del parametro

Sia X uno spazio metrico completo e sia Y uno spazio metrico (topologico). Sia $T\colon X \times Y \to X$ tale che $\exists \alpha < 1$ tale che $\forall y \in Y$, $T(-, y)$ è in $\mathrm{Lip}_\alpha(X)$. (α deve essere uniforme rispetto a y). Sia $y_0 \in Y$ tale che $\forall x \in X$, $T(x, y)$ sia continua in y_0. Allora f è continua in y_0.

Proof

Vogliamo mostrare che $d(f(y), f(y_0)) \to 0$ se $y \to y_0$. Vogliamo stimare $d(f(y), f(y_0))$ con $f(y) = T(f(y), y)$. Sia $x = f(y)$ e $f(y_0) = x_0$. Allora

$$d(f(y), f(y_0)) = d(T(x, y), T(x_0, y_0))$$

Usando la disuguaglianza triangolare

$$\begin{aligned} d(T(f(y_0), y_0)) &\leq d(T(f(y), y), T(f(y_0), y)) + d(T(f(y_0), y), T(f(y_0), y_0)) \\ &\leq \alpha d(f(y), f(y_0)) + d(T(f(y_0), y), T(f(y_0), y_0)) \\ (1 - \alpha) d(f(y), f(y_0)) &\leq d(T(f(y_0), y), T(f(y_0), y_0)) \to 0 \end{aligned}$$

La costante è positiva e indipendente da y. La funzione $T(f(y_0), -)$ è continua in y.

Lemma Sugli spazi normati

$$||y|| - ||x||| \leq ||y - x||$$

Proof

Sia $y = x + (y - x)$. Allora

$$||y|| = ||x + (y - x)|| \leq ||x|| + ||y - x||$$

Scambiando i ruoli di x e y si ottiene la proposizione.

Ciò mostra che la norma è lipschitz continua.

Ogni spazio normato è uno spazio metrico, ma non il viceversa.

Definizione Spazio di Banach

Uno *spazio di Banach* è uno spazio normato completo rispetto alla norma.

Definizione Equivalenza di norme

Diciamo che due norme $|| \cdot ||, | \cdot |$ sono equivalenti se $\exists 0 < \alpha \leq \beta$ tale che

$$\alpha |x| \leq ||x|| \leq \beta |x|, \quad \forall x \in X$$

Questa è una relazione di equivalenza.

Teorema Equivalenza di norme reali
Tutte le norme in $\mathbb{R}^n$ sono equivalenti.

Proof
Basta mostrare che una norma $||\cdot||$ questa è equivalente alla $||\cdot||_2$. Dobbiamo trovare α, β tale che
$$\alpha||x||_2 \leq ||x|| \leq \beta||x||_2, \quad \forall x \in \mathbb{R}^n$$
Consideriamo la base canonica $\{e_1, \cdots, e_n\}$ che è finito-dimensionale e quindi $x = (x_1, \cdots, x_n)$.
$$\begin{aligned}
||x|| &= \left|\left|\sum_{i=1}^{n} x_i \cdot e_i\right|\right| \leq \sum_{i=1}^{n} |x_1| \cdot ||e_i|| \\
&\leq \left(n \max \bigcup_{i=1}^{n} \{||e_i||\}\right) \left(\sum_{i=1}^{n} |x_i|\right) \\
&\leq \left(n \max \bigcup_{i=1}^{n} \{||e_i||\}\right) ||x||_1 \leq \underbrace{\left(n \max \bigcup_{i=1}^{n} \{||e_i||\}\right)}_{\beta} ||x||_2
\end{aligned}$$
Questo ci dice anche che $||\cdot||$ è continua rispetto alla topologia indotta da $||\cdot||_2$, e pure lipschitziana. Dobbiamo ora dimostrare l'altra metà della disuguaglianza e trovare α. Poniamo α in funzione dei vettori
$$\alpha = \inf_{||x||_2=1} ||x||$$
Mostriamo che $\alpha > 0$. Una volta fatto questo, possiam ottenere che la definizione di α dice che $||x||_2 = 1 \implies ||x|| \geq \alpha$. Voglio dimostrare che $||x|| \geq \alpha||x||_2$ per tutte le $x \in \mathbb{R}^n$. Se $x = 0$ la disuguaglianza è soddisfstta. Altrimnti, normalizziamo $z = x/||x||_2$. Usando l'omogeneità assoluta
$$||z||_2 = \left|\left|\frac{x}{||x||_2}\right|\right| = \frac{1}{||x||_2}||x||_2 \implies ||z||_2 = 1$$
Quindi $||z|| \geq \alpha$ and furthermore
$$\left|\left|\frac{x}{||x||_2}\right|\right| \geq \alpha \implies ||x|| \geq \alpha||x||_2$$
We now need to show that α is positive. Siccome le normi sono non-negative, alla peggio sono nulle. In realtà α è un minimo
$$\alpha = \min_{||x||_2=1} ||x||$$
since the norm is continuous with the respect to the topology induced by $||\cdot||_2$. The set over which we are taking the minimum is clearly closed and bounded. Siccome siamo nella topologia reale con norma euclidea è quindi anche compatto. Per Weierstrass, α è un minimo. Quindi deve essere $\alpha > 0$. Se fosse $\alpha = 0$ allora esisterebbe $\hat{x}$ tale che $||\hat{x}||_2 = 1$ e $||\hat{x}|| = 0$, che è assurdo lightning.

2 Operatori lineari fra spazi vettoriali

Lo spazio $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) \cong \mathrm{Mat}_{m\times n}(\mathbb{R})$.

Studiamo la continuità degli operatori lineari in spazi normati.

Proposition

Ogni $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ è continua rispetto alla topologia indotta dalla norma (qualsiasi visto che sono equivalenti in $\mathbb{R}^n$).

Proof

1. Mostriamo che la continuità dell'operatore in un singolo punto, come l'origine, implica la continuità di A in tutto $\mathbb{R}^n$. Questo è un fatto generale. Abbiamo quindi che se $\{x_n\} \to 0$ allora $\{Ax_n\} \to A0 = 0$. La continuità generale è data dal fatto che se $\{y_n\} \to x$ allora $\{Ay_n\} \to Ax$. Ma $\{Ay_n\} \to Ax$ se e solo se $\{Ay_m - Ax\} \to 0$ cioè $\{A(y_m - x)\} \to 0$ e per linearità $\{y_m - x\}$ è una successione che tende a zero, quindi A è continuo ovunque. Inoltre, per la continuità uniforme possiamo msotrare che $\forall \varepsilon > 0, \exists \delta > 0$ tale che
$$||y - x|| < \delta \implies ||Ay - Ax|| < \varepsilon$$
Ma $||Ay - Ax|| = ||A(x-y)||$. Poniamo quindi $z = y - x$. Dobbiamo mostrare che se $||z|| < \delta$ allora $||Az|| < \varepsilon$. Ma questa è la continuità nell'origine che stiamo presupponendo.
2. Mostriamo ora la continuità nell'origine. Usiamo il fatto che $\mathbb{R}^n$ ha dimensione finita. Sia $x = (x_1, \cdots, x_n)$ secondo la base canonica $(e_1, \cdots, e_n)$. Calcolo usando la disuguaglianza triangolare
$$\begin{aligned} ||Ax|| &= \left|\left|\sum_{i=1}^{n} x_i A e_i\right|\right| \leq \sum_{i=1}^{n} |x_i| \cdot ||Ae_i|| \\ &\leq C \sum_{i=1}^{n} |x_i|_1, \quad C = \max \bigcup_{i=1}^{n} \{||Ae_i||\} \end{aligned}$$
Ciò dimostra quindi che la funzione è Lipschitz-continua.

Nel passo secondo abbiamo usato il fatto che lo spazio fosse finitamente generato (il dominio). In generale, con $A \colon X \to Y$ è un operatore lineare fra spazi normati qualunque, on è detto che A sia continuo.

Esempio Controesempio

Consideriamo lo spazio delle funzioni $f \colon \mathbb{R} \to \mathbb{R}$ limitate, in $\mathcal{C}^1$ e con derivata limitata $\mathcal{BC}^1(\mathbb{R})$. Come secondo spazio prendiamo delle funzioni continue e limitate $f \colon \mathbb{R} \to \mathbb{R}$, $\mathcal{BC}(\mathbb{R})$. Consideriamo quindi l'operatore della derivata $\mathcal{BC}^1(\mathbb{R}) \to \mathcal{BC}(\mathbb{R})$. Siccome questi non sono spazi finitamente generati, dobbiamo scegliere delle norme. Scegliamo come norma sia nel dominio che nel codominio $|| \cdot ||_\infty$, che ha senso siccome le funzioni sono limitate. Mostriamo quindi che l'operatore lineare non è continuo. Scegliamo l'origine. Vogliamo quindi trovare $\{f_n\} \to 0$ ma tale che $\{f'_n\}$ non tende a zero. Per farlo prendiamo una funzione oscillante che si schiaccia sull'ascisse, e quindi la sua derivata non si schiaccia come la funzione. Prendiamo
$$f_n(x) = \frac{1}{n} \sin(nx)$$
Abbiamo la norma
$$||f||_\infty = \sup_{\mathbb{R}} \frac{1}{n} |\sin(nx)| = \frac{1}{n}$$
Mentre la norma della derivata
$$||f'||_\infty = \sup_{\mathbb{R}} |\cos(nx)| = 1$$

Andiamo a definire una norma speciale su questo spazio, la norma operatoriale.

Definizione Norma operatoriale

$$||A||_* = \sup_{||x||\leq 1} ||Ax||$$

Proposition

Norma operatoriale è una norma.

Mostriamo che la norma è ben definita, e che questo è un numero reale. Infatti, $||A||_* < +\infty$, siccome l'insieme del supremum è chiuso e limitato e, quindi, compatto, e la funzione è continua allora il supremum è un massimo. Mostriamo che il massimo si ottiene sulla frontiera della bolla.

Proof

Mostriamo le due disuguaglianze. Il fatto che $||A||_* \geq \max ||Ax||$ è banale, infatti $\{x \,|\, ||x|| = 1\}$ è un sottoinsieme di $\{x \,|\, ||x|| \leq 1\}$. D'altra parte se il massimo è ottenuto per $x = 0$ allora il max è 0 e quindi $Ax = 0$ sempre, e la disuguaglianza è banalmente soddisfatta. Supponiamo ora che il massimo sia ottenuto in un punto non nullo $\hat{x}$, possiamo normalizzare e ottenere norma unitaria.

$$||A\hat{x}|| = \left|\left|A\left(||\hat{x}||\frac{\hat{x}}{||\hat{x}||}\right)\right|\right| = \left|\left|\,||\hat{x}||A\left(\frac{\hat{x}}{||\hat{x}||}\right)\right|\right| = ||\hat{x}||\left|\left|A\left(\frac{\hat{x}}{||\hat{x}||}\right)\right|\right| \leq \left|\left|A\left(\frac{\hat{x}}{||\hat{x}||}\right)\right|\right|$$

$$||A||_* = \max_{||x||=1||Ax||} = |A\hat{x}| \leq \left|\left|A\frac{\hat{x}}{||\hat{x}||}\right|\right| \leq \max_{||x||=1} ||Ax||$$

Quindi il valore può essere calcolato solo sulla buccia in quanto lì viene raggiunto il massimo.

Proposition Stima fondamentale

Per ogni $x \in \mathbb{R}^n$ vale

$$||Ax||_* \leq ||A|| \cdot ||x||$$

Proof

Se $||x|| = 1$ vale $||Ax|| \leq ||A||$ perché per la proprietà precedente,

$$||A|| = \max_{||x||=1} ||Ax||$$

Se $x = 0$, la disuguaglianza vale. Altrimenti, normalizziamo x per ritrovarci sulla frontiera.

$$||A||_* \geq \left|\left|A\left(\frac{x}{||x||}\right)\right|\right| = \frac{1}{||x||}||Ax||$$

moltiplicando entrmambi i membri epr $||x||$ si ottiene la tesi.

In realtà questa costante è la migliore.

Proposition

Se $\exists \alpha \in \mathbb{R}$ tale che $||Ax|| \leq \alpha||x||$ per tutte le $x \in \mathbb{R}^n$, allora

$$||A||_* \leq \alpha$$

In realtà questo risultato vale anche se supponiamo che $||Ax|| \leq \alpha||x||$ solo per x tale che $||x|| \leq 1$ oppure tale che $||x|| = \varepsilon$ per qualche $\varepsilon > 0$.

Proof

Supponiamo di avere una stima del tipo $||Ax|| \leq \alpha||x||$ per tutte le x. Sappiamo che

$$||A||_* = \max_{||x||=1} ||Ax|| \leq \alpha \max_{||x||\leq 1} ||x|| = \alpha$$

che è quindi chiaramente 1.

La medesima dimostrazione funziona supponendo che $||Ax|| \leq \alpha||x||$ per tutte le $||x|| \leq 1$. Se invece sappiamo che $||Ax|| \leq \alpha||x||$ solo per gli x tale che $||x|| = \varepsilon$, allora è sufficiente normalizzare (per esercizio).

Proof La norma operatoriale è una norma

1. *annullamento:* Supponiamo che $||A||_* = 0$. Devo mostrare che $A = 0$. Siccome la norma è nulla,
$$\max_{||x||\leq 1} ||Ax|| = 0 \implies ||Ax|| = 0, \quad \forall x \mid ||x|| \leq 1$$
e quindi anche per tutti le altre x visto che possiamo normalizzare. Quindi, $Ax = 0$.
2. *positiva omogeneità:*
$$||\lambda A||_* = \max_{||x||=1} ||\lambda Ax|| = \max_{||x||=1} |\lambda||Ax| = |\lambda| \max_{||x||=1} ||Ax|| = |\lambda|||A||_*$$
3. *disuguaglianza triangolare:*
$$\begin{aligned} ||(A+B)x|| &= ||Ax + Bx|| \leq ||Ax|| + ||Bx|| \\ &\leq ||A||_*||x|| + ||B||_*||x| = ||x|| \left(||A||_* + ||B||_*\right) \end{aligned}$$
Per la proposizion precedente, $||A + B||_*$ è la più piccola costante per cui vale una disuguaglianza di questo tipo.

La norma operatoriale è scelta tale precisamente per la stima.

Teorema

$(\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m), || \cdot ||)$ è uno spazio di Banach.

Proof

Siccome questo spazio è finito dimensionale è isomorfismo allo spazio $\mathbb{R}^{m\times n}$ che è completo per esempio rispetto alla norma seconda. Tuttavia, dimostramolo con le successioni di Cauchy. Cosideriamo quindi una successione di operatori lineari. Per tutte le $\varepsilon > 0$ esiste N_ε tale che $\forall m, n > N_\varepsilon$ tale che

$$||A_m - A_n|| < \varepsilon$$

cioè per tutte le x

$$||(A_n - A_n)x|| \leq \varepsilon(x)$$

Questo vuole dire che per x fissato la successione $\{A_n x\}$ è una successione di Cauchy in $\mathbb{R}^n$. Siccome $\mathbb{R}^n$ è completo, la successione converge. Chiamiamo il limite di tale successione Ax. Verifichiamo che in questo modo abbiamo definito un operatore lineare $x \to Ax$. Sappiamo che A_n è lineare per ogni n, quindi $A_n(x+y) = A_n x + A_n y$. Sappiamo che $A(x+y)$ tende ad $Ax+Ay$ e quindi l'espressione sopra tende ad $Ax + Ay$. Analogamente per l'omogeneità. Mostriamo ora che $\{A_n\}$ effettivamente converge ad A, quindi $||A_n - A|| \to 0$. Sappiamo che $\{A_n\}$ è una successione di Cauchy. Quindi $\forall\varepsilon > 0$ esiste N_ε tale che $\forall n, m > N_\varepsilon$

$$||A_m x - A_n x| \leq \varepsilon||x||$$

Ma $A_n x \to Ax$ per $n \to \infty$. passando al limite si ha che per tutte le $\varepsilon > 0$ esiste N_ε tale che $\forall n > N_\varepsilon$,

$$||A_n x - Ax|| \leq \varepsilon |x|$$

cioè abbiamo trovato che $||A_n - A|| - \varepsilon$.

Notiamo che abbiamo sfruttato solo la completezza di $\mathbb{R}^n$.

Esercizio

Sia $A\colon \mathbb{R}^2 \to \mathbb{R}^3$ dato da

$$A = \begin{pmatrix} 3 & -2 \\ 0 & 1 \\ 1 & 3 \end{pmatrix}$$

Prendiamo la norma infinito in ambo gli spazi. Il massimo è dato dal valore sulla frontiera della bolla.

$$||A||_* = \max_{\max\{|x|,|y|\}=1} \max\{|3x - 2y|, |y|, |x + 3y|\}$$

Quindi $|3x - 2x| \leq 3|x| + 2|y|$ edeve essere $\max\{||x||, ||y||\} = 1$. In particolare i due moduli sono minori o uguali ad uno. $|3x - 2y| \leq 3 + 2 = 5$ e il valore 5 è assunto per $x = 1$ e $y = -1$ (oppure cil contrario). Quindi $||A|| = 5$.
Consideriamo invece ora la norma 1 quindi

$$||A|| = \max_{|x|+|y|=1} (|3x - 2y| + |y| + |x + 3y|)$$

Allora

$$|3x - 2y| + |y| + |x + 3y| \leq 3|x| + 2|y| + |y| + |x| + 3|y| = 4|x| + 6|y|$$

Siccome $|x| + |y| = 1$ quest'ultima è un interpolazione lineare tipo $ta + (1 - t)b$. In questo caso abbiamo l'espressione $4|x| + 6|y|$ che è sempre minore o uguale di 6. Quindi abbiamo trovato che $|3x - 2y| + |y| + |x + 3y| \leq 6$. Tale valore può essere raggiunto per esempio scegliendo $x = 0$ e $y = 1$.

Consideriamo ora operatori lineari invertibili, quindi fra spazi con la medesima dimensione.

Teorema

$\mathrm{GL}_n(\mathbb{R})$ è un aperto in $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ dove consideriamo lo spazio topologico indotto dalla metrica indotta dalla norma associata allo spazio. Inoltre, la mappa dell'inversione è continua.

Proof

Visto che il determinante è non nullo

$$\mathrm{GL}_n(\mathbb{R}) = \det^{-1}(\mathbb{R} \backslash \{0\})$$

ma visto che è la preimmagine di un aperto tramite il determinante, che è continuo.
Dimostriamo tuttavia la proposizione con una argomento che può funzionare anche nel caso infinito dimensionale.

1. *identità è un punto interno dell'insieme:* mostriamo che esiste δ tale che $||A - \mathbf{1}|| < \delta$ implica che $A \in \mathrm{GL}(\mathbb{R}^n)$. Stiamo usando la norma operatoriale. Posiamo $B = \mathbf{1} - A$ e quindi $A^{-1} = (\mathbf{1} - B)^{-1}$ che possiamo scrivere come la serie geometrica di B^k. Ovviamente dobbiamo usare la convergenza della somma parziale della serie secondo la norma dei vettori. (Purché la serie converga). Infatti, se la serie converge usiamo la proprietà distributiva per

mostrare che è veramente l'inverso di $\mathbf{1}-B$

$$\begin{aligned}(\mathbf{1}-B)\sum_{k=0}^{\infty}B^k &= \sum_{k=0}^{\infty}B^k - \sum_{k=0}^{\infty}B^{k+1}\\ &= \sum_{k=0}^{\infty}B^k - \sum_{k=1}^{\infty}B^k = B^0 = \mathbf{1}\end{aligned}$$

Quindi la stessa cosa che vale per i numeri, vale anche con gli operatori (negli spazi vettoriali in generale). Dobbiamo tuttavia studiare quando la serie converge. Usando il criterio di Cauchy, troviamo che la serie converge se e solo se $\forall\varepsilon > 0, \exists N_\varepsilon \in \mathbb{N}$ tale che $\forall p \in \mathbb{N}$

$$\left|\left|\sum_{k=m}^{m+p}B^k\right|\right| \leq \varepsilon$$

Mediante la stima otteniamo una serie numerica.

$$\left|\left|\sum_{k=m}^{m+p}B^k\right|\right| \leq \sum_{k=m}^{m+p}||B^k||$$

Preferiremmo lavorare la potenza di una norma piuttosto che la norma di una potenza, ma coincidono?
Calcoliamo quindi $||B\circ A||_*$ con A, B operatori lineari, e usando la norma operatoriale. Siccome abbiamo la norma operatoriale vale la stima fondamentale

$$||(B\circ A)x|| = ||B(Ax)|| \leq ||B||_*||Ax|| \leq ||B||_*||A||_*||x||$$

Siccome la norma operatoriale è la più piccola costante per cui vale una stima del tipo dato. Quindi $||B\circ A||_* \leq ||B||_*||A||_*$. Quindi in particolare $||B^k||_* \leq ||B||_*{}^k$. Tornando a prima otteniamo

$$\left|\left|\sum_{k=m}^{m+p}B^k\right|\right|_* \leq \sum_{k=m}^{m+p}||B^k||_* \leq \sum_{k=m}^{m+p}\left(||B||_*\right)^k \leq \varepsilon$$

pur di considerare $||B||_* < 1$.

2. Sia $A_0 \in \mathrm{GL}_n(\mathbb{R})$. Vogliamo mostrare che $\exists\delta > 0$ tale che

$$||A - A_0||_* < \delta \implies A \in \mathrm{GL}_n(\mathbb{R})$$

Calcoliamo

$$A = A_0 - (A_0 - A) = A_0(\mathbf{1}A_0^{-1}(A_0 - A)) = A_0(\mathbf{1}-B), \quad B = \mathbf{1}-B$$

per il punto precedente $\mathbf{1}-B$ è invertibile purché $||B||_* < 1$. Studiamo quando ciò avviene

$$||B||_* = ||A_0^{-1}(A_0 - A)||_* \leq ||A_0^{-1}||\cdot||A_0 - A|| < 1 \text{ se } ||A_0 - A||_* < \frac{1}{||A_0^{-1}||}$$

la divisione la possiamo fare in quanto la funzione nulla, che ha norma nulla, non è invertibile quindi non siamo in quel caso. Quindi, scegliendo $\delta \leq 1/||A_0^{-1}||$ si ha che A è invertibile, quindi è un aperto. Cioè inq uesto caso $||B|| < 1$ quindi $\mathbf{1}-B$ è invertibile e visto che abbiamo scritto A come prodoto di operatori invertibili, è anch'esso invertibile.

Mostriamo ora che l'inversa è continua. Da $||B\circ A||_* \leq ||B||_*||A||_*$ da qui segue che la composizione di questi lineari è continua. Siano $B_n \to B$ e $A_n \to A$ quindi $B_n \circ A_n \to B\circ A$. Mostriamo che la mappa $A \to A^{-1}$ è continua, cioè che $A_n \to A \implies A_n^{-1} \to A^{-1}$. Abbiamo

$$A = A_0(\mathbf{1} - A_0^{-1}(A_0 - A)) = A_0(\mathbf{1}-B), \quad A^{-1} = (\mathbf{1}-B)^{-1}A_0^{-1}$$

Vogliamo stimare $||A^{-1} - A_0^{-1}||_*$.

$$A^{-1} - A_0^{-1} = (\mathbf{1} - B)^{-1} A_0^{-1} - A_0^{-1} = \left[(\mathbf{1} - B)^{-1} - \mathbf{1}\right] A_0^{-1} = \left[\sum_{k=0}^{\infty} B^k - \mathbf{1}\right] A_0^{-1} = \left[\sum_{k=1}^{\infty} B^k\right] A_0^{-1}$$

$$||A^{-1} - A_0^{-1}||_* = \left|\left|\left(\sum_{k=1}^{\infty} B^k\right) \circ A_0^{-1}\right|\right| \leq \left(\sum_{k=1}^{\infty} (||B||_*)^k\right) ||A_0^{-1}|| = \frac{||B||_*}{1 - ||B||_*} ||A_0^{-1}||$$

pur di prendere $||B||_* < 1$. Se $A \to A_0 \implies ||B||_* \to 0$ allora $A^{-1} - A_0^{-1} \to 0$.

3 Calcoli differenziale multivariabile

Definizione Limite multivariabile

Sia $f\colon \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^m$ e sia a un punto di accimulazione per Ω. Allora

$$\lim_{x \to a} f(x) = b \in \mathbb{R}^m$$

se $\forall \varepsilon > 0, \exists \delta > 0$ tale che $\forall x \in \Omega$

$$0 < ||x - a||_1 < \delta \implies ||f(x) - b||_2 < \varepsilon$$

dove le norme sono generiche.

Il concetto di limite ad infinito vuol dire uscire dai compatti.

Definizione Limite a infinito multivariabile

Sia $f\colon \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^m$ con Ω illimitato. Allora

$$\lim_{x \to \infty} f(x) = b \in \mathbb{R}^m$$

se $\forall \varepsilon > 0, \exists M > 0$ tale che $\forall x \in \Omega$

$$||x|| > M \implies ||f(x) - b|| < \varepsilon$$

e per il codominio

Definizione Limite infinito multivariabile

Sia $f\colon \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^m$ e a punto di accumulazione. Allora

$$\lim_{x \to a} f(x) = \infty$$

se $\forall K > 0, \exists \delta > 0$ tale che per tutte le $x \in \Omega$

$$0 < ||x - a|| < \delta \implies ||f(x)|| > K$$

Definizione Cammino

Un *cammino* è un applicazione $\varphi : (0, 1) \to \mathbb{R}^n$ continua.

Sia $\Phi_\Omega(a)$ l'insieme dei cammini con sostegni in Ω e tale che $\varphi(t) \to a$ per $t \to 0^+$ e $\varphi(t) \neq$ per tutte le t.

Teorema

$$\lim_{x \to a} f(x) = b \iff \lim_{t \to 0^+} f(\varphi(t)) = b$$

per tutte le $\varphi \in \Phi_\Omega(a)$.

Definizione Derivata direzionale

TODO

$$D_v f(a) = \lim_{t \to 0} \frac{f(a + vt) - f(a)}{t}$$

La derivata parziale è come prendere una fetta della funzione multidimensionale, nella direzione data dal vettore direzionale, e calcolare la verivata ad una singola variabile.

Le derivate parziali lungo i versori della base canonica si chiamano derivate parziali.

Esempio

Sia

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

Stabiliamo se le derivate direzionali di f esistono nell'origine e calcoliamo. Questa funzione è costante lungo gli assi cartesiani e vale zero. Quindi, le derivate parziali sono nulle. Sia v un qualsiasi versore. Possiamo già escludere i casi con $v_i = 0$ in quanto sono le derivate parziali. Calcoliam quindi

$$\begin{aligned} \lim_{t\to 0} \frac{f(0+tv_1, 0+tv_2) - f(0,0)}{t} &= \lim_{t\to 0} \frac{1}{t} \frac{tv_1 t^2 v_2^2}{t^2 v_1^2 + t^2 v_2^2} \\ &= \lim_{t\to 0} \frac{t^3 v_1 v_2^2}{t^3(v_1^2 + t^2 v_2^4)} \\ &= \lim_{t\to 0} \frac{v_1 v_2^2}{v_1^2 + t^2 v_2^4} \\ &= \frac{v_1 v_2^2}{v_1^2} = \frac{v_2^2}{v_1} \end{aligned}$$

Quindi f ammette tutte le derivate direzionali. Controlliamo il limite di questa funzione nell'origine, per controllare se è quindi continua in quel punto.

$$\frac{\rho^3 |\cos\theta| \sin^2\theta}{\rho^2(\cos^2\theta + \rho^2 \sin^2\theta)} \leq \frac{\rho}{\cos^2\theta + \rho^2 \sin^4\rho}$$

Consideriamo

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

Quindi abbimo trovato un cammino lungo il qualr la funzione non tende a zero. Avremmo potuto semplicemente considerare la stima

$$\frac{|xy^2|}{x^2+y^4} \leq \frac{1}{2}$$

L'esistenza delle derivate parziali non implica la differenziabilità.

Nelle funzioni di una variabili l'esistenza del rapporto incrementale e l'approssimazione lineare local di una funzione coincidono e sono equivalenti alla differenziabilità, ma in più variaibli non coincidono.

Definizione

Sia $f\colon \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ con Ω aperto e $a \in \Omega$. Diremo che f è differenziabile in a se esiste una forma lineare (funzionale lineare), o equivalentemente se esiste $\lambda \in \mathbb{R}^n$ tale che

$$f(a+h) = f(a) + \lambda\lambda + o(||h||)$$

per $h \to 0$. Tale forma lieare è detta differenziale di f in a, denotata $\mathrm{d}f(a)$, che quindi manda h in $h\lambda$.

(Teorema di Riesz, infatti il fatto che la derivata direzionale è data da $\nabla f(a) \cdot v$). Se tale λ esiste, è unico. Supponiamo che λ, μ soddisfino l'uguaglianza. Sottraendo membro a membro otteniamo

$$(\lambda - \mu)h = o(||h||)$$

per $h \to 0$. Quindi vogliamo

$$\lim_{h\to 0} \frac{(\lambda-\mu)h}{||h||} = 0$$

ma

$$(\lambda-\mu)\frac{h}{||h||}$$

non tende a zero per $h \to 0$. Per mostrarlo scegliamo il cammino $\varphi(t) = t(\mu - \nu)$ e

$$(\lambda-\mu)\frac{t(\lambda-\mu)}{|t|\cdot||\lambda-\mu||} = \frac{t}{|t|}||\lambda-\mu||$$

che nont ende a zero perché $\lambda \neq \mu$. Quindi non esistono $\lambda \neq \mu$ tale che valga l'uguaglianza.

Teorema

Se f è differenziabile in a, allora

1. f è continua in a;
2. $\forall v \in \mathbb{R}^n$, esiste la derivata direzionale $D_v f(a) = \lambda \cdot v$;
3. $\lambda = \nabla f(a)$

Proof

1. Per ipotesi esiste $\lambda \in \mathbb{R}^n$ tale che
$$f(a+h) - f(a) = \lambda \cdot h + \varepsilon(h)$$
con $\varepsilon(h) = o(||h||)$. Quindi,
$$\begin{aligned} |f(a+h) - f(a)| &= |\lambda \cdot h + \varepsilon(h)| \\ &= |\lambda h| + |\varepsilon(h)| < ||\lambda|| ||h|| + ||h|| \frac{|\varepsilon(h)|}{||h||} \\ &\leq ||h|| \left[||\lambda|| + \frac{|\varepsilon(h)|}{||h||} \right] \to 0 \end{aligned}$$
Abbiamo usato la disuguaglianza di Cauchy-Schwarz.
2. Calcoliamo
$$\begin{aligned} \frac{f(a+tv) - f(a)}{t} &= \frac{\lambda \cdot tv + \varepsilon(tv)}{t} \\ &= \lambda v + \frac{\varepsilon(tv)}{t} \\ &= \lambda v + \frac{\varepsilon(tv)}{t||v||} \\ &= \lambda v + \frac{|t|}{t}\frac{\varepsilon(tv)}{||tv||} \\ &\to \lambda v \end{aligned}$$
3. $$\frac{\partial f}{\partial x_j}(a) = D_v f(a) = \lambda e_j = \lambda_j$$
quindi $\lambda = \nabla f(a)$.

Per sabilire se una funzione f è differenziabile in a, basta controllare se le derivate parziali esistono.

Corollario Formula del gradiente

Se f è differenziabile in a, allora $D_v f(a) = \nabla f(a) \cdot v$.

Esempio

Studia se

$$f(x, y) = x^5 \sqrt{|y|}$$

è differenziabile nei punti del tipo $(x_0, 0)$ lungo l'asse x. Nel punto zero potremo avere dei problemi per la derivata parziale di y, quindi dovremo usare la definizione e non le regole delle derivate. Siccome $f(x, 0) = 0$ la derivata nei punti della forma $(x_0, 0)$ è 0. Calcoliamo

$$\begin{aligned}
\frac{\partial f}{\partial y}(x_0, 0) &= \lim_{y \to 0} \frac{f(x_0, y) - f(y_0, 0)}{y} \\
&= \lim_{y \to 0} \frac{x_0^5 \sqrt{||y||} - 0}{y} \\
&= \lim_{y \to 0} x_0^5 \frac{\sqrt{|y|}}{y} \\
&= \lim_{y \to 0} x_0^5 \operatorname{sgn} y |y|^{-1/2} = \begin{cases} 0 & x_0 = 0 \\ \nexists & x_0 \neq 0 \end{cases}
\end{aligned}$$

Ma quindi f non è differenziabile nei punti $(x_0, 0)$ se $x_0 \neq 0$. Invece nel punto dell'origine abbiamo $\nabla f(0, 0) = (0, 0)$

$$\lim_{(h,k) \to (0,0)} \frac{f(h, k) - f(0, 0) - \nabla f(0, 0) \cdot (h, k)}{\sqrt{h^2 + k^2}} = \lim_{(h,k) \to (0,0)} \frac{h^5 \sqrt{|k|}}{\sqrt{h^2 + k^2}}$$

Siccome

$$\frac{|ab|}{a^2 + b^2} \leq \frac{1}{2}$$

Abbiamo

$$\left| \frac{h^5 \sqrt{|k|}}{\sqrt{h^2 + k^2}} \right| = \left(\frac{|h| \cdot |k|}{h^2 + k^2} \right)^{1/2} |h|^{9/2} \leq \frac{1}{\sqrt{2}} |h|^{9/2}$$

e quindi tende a zero. Allora f è differenziabile nell'origine.

Se f è differenziabile in a possiamo definire l'iperpiano tangente

$$f(a) + \nabla f(a) \cdot (x - a)$$

Teorema Teorema del differenziale totale

Sia $f \colon \Omega \to \mathbb{R}$, con $\Omega \in \mathbb{R}^n$ aperto, $x_0 \in \Omega$. Se $\exists r > 0$ tale che $B_r(x_0) \subseteq \Omega$ e tutte le derivate parziali di f esistono in $B_r(x_0)$ e sono continue in x_0, allora f è differenziabile in x_0.

È una condizione sufficiente.

Proof

Facciamo il caso $n = 2$. Per ipotesi sappiamo che $\nabla f(x_0)$ esiste. Dobbiamo msotrare che

$$\lim_{(x,y) \to (x_0,y_0)} \frac{f(x, y) - f(x_0, y_0) - \nabla f(x) \cdot (x - x_0, y - y_0)}{\sqrt{(x - x_0)^2 + (y - y_0)^2}} = 0$$

Calcoliamo l'incremento $f(x,y) - f(x_0,y_0)$. Piuttosto che congiungere i due punti mediante un segmento, facciamo lo stesso percorso con due segmenti nella direzione degli assi, quindi ho una variabile sola al posto di due, due volte. Quindi passiamo dal punto intermedio (x, y_0). Abbiamo allora

$$f(x,y) - f(x_0,y_0) = [f(x,y) - f(x,y_0)] + [f(x,y_0) - f(x_0,y)]$$

Possiamo applicare il teorema di Lagrange

$$f(x,y) - f(x,y_0) = \frac{\partial f}{\partial y}(x,\xi)(y - y_0)$$

con $\xi \in [y, y_0]$, visto che la derivata è parziale. Considerando allo stesso modo $f(x,y_0)$ e applicando Lagrange si ottiene

$$f(x,y_0) - f(x_0,y_0) = \frac{\partial f}{\partial x}(\mu,y_0)(x - x_0)$$

dove $\eta \in [x, x_0]$. Sostituendo troviamo

$$\frac{f(x,y) - f(x_0,y_0) - \frac{\partial f}{\partial x}(x_0,y_0)(x-x_0) - \frac{\partial f}{\partial y}(x_0,y_0)(y-y_0)}{\sqrt{(x-x_0)^2 + (y-y_0)^2}}$$
$$= \left| \frac{\left[\frac{\partial f}{\partial x}(\eta,y_0) - \frac{\partial f}{\partial x}(x_0,y_0)\right](x-x_0) + \left[\frac{\partial f}{\partial y}(x,\xi) - \frac{\partial f}{\partial y}(x_0,y_0)\right](x-x_0)}{\sqrt{(x-x_0)^2 + (y-y_0)^2}} \right|$$
$$= \left|\frac{\partial f}{\partial x}(\eta,y_0) - \frac{\partial f}{\partial x}(x_0,y_0)\right| \frac{|x-x_0|}{\sqrt{(x-x_0)^2 + (y-y_0)^2}} + \left|\frac{\partial f}{\partial y}(x,\xi) - \frac{\partial f}{\partial y}(x_0,y_0)\right| \frac{|y-y_0|}{\sqrt{(x-x_0)^2 + (y-y_0)^2}}$$
$$\leq \left|\frac{\partial f}{\partial x}(\eta,y_0) - \frac{\partial f}{\partial x}(x_0,y_0)\right| + \left|\frac{\partial f}{\partial y}(x,\xi) - \frac{\partial f}{\partial y}(x_0,y_0)\right|$$

I due termini con le radici a denominatori sono minori di 1. Per il teorema del confronto, se $(x,y) \to (x_0,y_0)$ implica che $\xi \to y_0$ e $\eta \to x_0$. Per la continuità della derivate parziali segue la tesi che il valore tende a zero.

Corollario

Se in Ω tutte le derivate parziali di f esistono e sono continue, allora f è differenziabile in Ω.

Definizione

Sia $f\colon \Omega \to \mathbb{R}$, con $\Omega \in \mathbb{R}^n$ aperto. Se le derivate parziali esistono e sono continue in Ω, diremo che f è differenziabile con continuità e scriviamo $f \in C^1(\Omega)$.

Per il teorema del differenziale totale le classi C sono annidate.

Esercizio

Studiare la differenziabilità di

$$f(x,y) = \ln(1 + xy) - x^2 y$$

. Il dominio massimalel è $\{(x,y) \,|\, xy > -1\}$. Troviamo le due derivate parziali

$$\frac{\partial f}{\partial x} = \frac{y}{1+xy} - 2xy, \quad \frac{\partial f}{\partial y} = \frac{x}{1+xy} - x^2$$

Siccome le derivate sono continue ed esistono ovunque nel dominio (in particolare nell'origine) la funzione è differenziabile.

Esercizio

Studiare continuità e differenziabilità nell'origine di

$$f(x,y) = \begin{cases} \sqrt{x^2+y^2}\left[1-e^{-\frac{x^2+y^2}{|x|}}\right] & x \neq 0 \\ 0 & x = 0 \end{cases}$$

Non conviene usare il tereoma del differenziale totale perché la funzione è a tratti e quindi dovremmo usare la definizione di derivata. Cominciamo con la continuità. Calcoliamo il limite $(x,y) \to (0,0)$ notando che $|f| \leq \sqrt{x^2+y^2} \to 0$. Quindi è continua nell'origine. Calcoliamo adesso le derivate direzionali. Sia $v = (v_1, v_2)$ un versore

$$\begin{aligned} \frac{f(tv)-f(0)}{t} &= \frac{1}{t}\left(|t|\left[1-e^{\frac{t^2}{|t|\cdot|v_1|}}\right]\right) \\ &= \operatorname{sgn} t\left[1-e^{-\frac{|t|}{|v_1|}}\right] \to 0 \end{aligned}$$

quando $v_1 \neq 0$. Nel caso $v_1 = 0$ abbiamo il vettore $(0, \pm 1)$ della direzione dell'asse y. In tal caso abbiamo

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

in quanto f è identicamente nulla in quell'asse. Ma f non è differenziabile in quanto il seguente non tende a zero

$$\left|\frac{f(x,y)-f(0,0)-\nabla f(0,0)\cdot(x,y)}{\sqrt{x^2+y^2}}\right| = 1-e^{-\frac{x^2+y^2}{|x|}}$$

e considerando l'esponente. Possiamo trovare un cammino per il quale il limite non tende a zero. Le rette non vanno bene, dobbiamo prendere un cammino che è tangente a $x = 0$. Prendiamo per esempio (t^2, t). L'esponente diventa $-(t^4+t^2)/t^2 \to -1$. Quindi tutto il limite tende a $1-e^{-1} \neq 0$.

La forma differenziabile $\mathrm{d}f(-)\colon \Omega \to \mathcal{L}(\mathbb{R}^n, \mathbb{R})$ e quindi associa ad ogni punto il funzionale lineare differenziale. Quindi

$$\mathrm{d}f(a)\colon \mathbb{R}^n \to \mathbb{R}$$

che associa $h \to \nabla f(a)\cdot h$. In due variabili per esempio

$$\mathrm{d}f(a)(x-a) = \nabla f(a)(x-a) = \frac{\partial f}{\partial x}(a)(x-e_1) + \frac{\partial f}{\partial x}(x-e_2)$$

la cui applicazuine può essere scritta con le proiezioni

$$\mathrm{d}f(a) = \frac{\partial f}{\partial x}(a)\,dx + \frac{\partial f}{\partial x}\,dy$$

Nota: $\nabla f(a)$ è il vettore che punta nella direzione di massima crescita locale. Infatti, $f(a+h)-f(a) = \nabla f(a)\cdot h + \varepsilon(h)$ con $\varepsilon(h) = o(||h||)$. Abbiamo

$$\begin{aligned} f(a+h)-f(a) &= ||\nabla f(a)||\cdot||h||\cos\theta + \varepsilon(h) \\ &= ||h||\left[||\nabla f(a)||\cos\theta + \frac{o(||h||)}{||h||}\right] \end{aligned}$$

dove θ è l'nagolo fra h e $\nabla f(a)$. Quindi ciò assume valore massimo quando $\theta = 0$, cioè quando la direzione è il verso di $\nabla f(a)$.

3.1 Funzioni a valori vettoriali

Consideriamo $f\colon \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^m$ con Ω aperto. Abbiamo quindi

$$f = \sum_{i=1}^{n} f_i e_i, \quad f_i\colon \Omega\mathbb{R}$$

Sia $a \in \Omega$ e sia $v \in \mathbb{R}^n$. Definiamo la derivata direzionale

$$D_v f(a) = \lim_{t\to 0} \frac{f(a+tv) - f(a)}{t} \in \mathbb{R}^m$$

e le componenti sono le derivate direzionali delle funzioni a valore reale. Posso definire un totale di mn derivate parziali

$$\frac{\partial f_i}{\partial x_j}(a)$$

Possiamo quindi definire la matrice Jacobiana come la matrice $J(a)_{i,j} = \frac{\partial f_i}{\partial x_j}(a)$. Quindi le righe della matrice sono i trasporti dei gradienti $\nabla f_j(a)$.

Definizione Differenziabilità

Una funzione $f\colon \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^m$ con Ω aperto è differenziabile in $a \in \Omega$ se $\exists L \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ tale che

$$f(a+h) - f(a) = L(h) + \varepsilon(h)$$

dove $\varepsilon(h) = o(||h||)$. Tale applicazione lineare di chiama differenziale.

Equivalentemente, f è differenziabile in a se esiste una matrice $\Lambda \in M_{m\times n}(\mathbb{R})$ tale che

$$f(a+h) - f(a) = \Lambda h + \varepsilon(h)$$

Equivalentemente f è differenziabile se e solo se f_i sono tutte differenziabili.

Teorema Teorema del differenziale totale

Vale l'analogo teorema: se tutte le derivate parziali delle componenti di f esistono in un intorno di a e sono continue in a allora f è differenziabile in a

Teorema

Se f è differenziabile in a allora f è continua in a, $\forall v \in \mathbb{R}^n$ tale che $||v|| = 1$,

$$\exists D_v f(a) = \Lambda v$$

e $\Lambda = J(a)$.

Cioè il differenziale (differenziale primo) L è rappresentato nella base canonica dalla matrice Jacobiana,

$$\mathrm{d}f_a h = Jf(a)h = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j}(a) h_j \right) e_i$$

3.2 Chain rule

Teorema Chain rule

Siano $f\colon \mathbb{R}^n \to \mathbb{R}^m$ e $g\colon \mathbb{R}^m \to \mathbb{R}^p$ funzioni tali che f è differenziabile nel punto a e g differenziabile in $f(a)$. Allora, $g \circ f$ è differenziabile in a e vale

$$\mathrm{d}(g \circ g)(a) = \mathrm{d}g(f(a)) \circ \mathrm{d}f(a)$$

O equivalentemente in forma matriciale

$$J(g \circ f)(a) = Jg(f(a))Jf(a)$$

Sia $h = g \circ f \colon \mathbb{R}^m \to \mathbb{R}^n$. Calcoliamo il seguente, che è il prodotto della i-esima riga di $Jg(f(a))$ per la j-esima colonna di $Jf(a)$

$$\frac{\partial h_i}{\partial x_i}(a) = \sum_{k=1}^{m} \frac{\partial g_i}{\partial y_k}(f(a)) \cdot \frac{\partial f_k}{\partial x_j}(a)$$

Esempio

Siano $f \colon \mathbb{R}^2 \to \mathbb{R}^3$ e $g \colon \mathbb{R}^3 \to \mathbb{R}^2$. Chiamiano x, y le coordinate di f e u, v, w quelle di g. Sia $h = g \circ f$ quindi $h(x, y) = (h_1(x, y), h_2(x, y))$. Calcoliamo per esempio

$$\frac{\partial f_2}{\partial x}(a) = \frac{\partial g_2}{\partial u}(f(a)) \cdot \frac{\partial f_1}{\partial x}(a) + \frac{\partial g_2}{\partial v}(f(a)) \cdot \frac{\partial f_2}{\partial x}(a) + \frac{\partial g_2}{\partial w}(f(a)) \cdot \frac{\partial f_3}{\partial x}(a)$$

Esempio

Siano $f \colon \Omega \subseteq \mathbb{R}^2 \to \mathbb{R}^2$ e $g \colon \mathbb{R}^2 \to \mathbb{R}$. Chiamiamo le variabili di f (ρ, θ) e di g dicimo (x, y). Definiamo $f(p, \theta) = (\rho \cos\theta, \rho \sin\theta)$. Sia $h = g \circ f$ quindi $h(\rho, \theta) = g(\rho\cos\theta, \rho\sin\theta)$. Calcoliamo

$$\frac{\partial h}{\partial \rho} = \frac{\partial g}{\partial x}(\rho\cos\theta, \rho\sin\theta)\cos\theta + \frac{\partial g}{\partial y}(\rho\cos\theta, \rho\sin\theta)\sin\theta$$

e

$$\frac{\partial h}{\partial \theta} = -\frac{\partial g}{\partial x}(\rho\cos\theta, \rho\sin\theta)\rho\sin\theta + \frac{\partial g}{\partial y}(\rho\cos\theta, \rho\sin\theta)\rho\cos\theta$$

Esercizio

Studiare la continuità e differenziabilità nell'origine di

$$f(x, y) = \begin{cases} \frac{|x|^{1/4}\ln(1+x^2+y^2)}{(x^2+y^2)^\alpha} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases}$$

al variare di α. Il dominio è tutto $\mathbb{R}^2$. Cominciamo studiando il limite nell'origine. In coordinate polari

$$\begin{aligned} F(\rho, \theta) &= \frac{\rho^{1/4}|\cos\theta|^{1/4}\ln(1+\rho^2)}{\rho^{2\alpha}} \leq \frac{\rho^{1/4}\ln(1+\rho^2)}{\rho^{2\alpha}} \\ &= \rho^{\frac{1}{4}-2\alpha}\ln(1+\rho^2) \leq \rho^{\rho^{\frac{1}{4}-2\alpha+2}} \end{aligned}$$

Quindi sicuramente tende a zero per $\alpha < 9/8$, dove f è continua. Per dimostrare che il bound è sharp supponiamo $\alpha \geq 9/8$ e consideriamo il cammino $x = 0$. Abbiamo

$$\begin{aligned} f(x, 0) &= \frac{|x|^{1/4}\ln(1+x^2)}{(x^2)^\alpha} \\ &= \frac{|x|^{1/4}\ln(1+x^2)}{|x|^{2\alpha}} \sim |x|^{\frac{1}{4}+2-2\alpha} \end{aligned}$$

Da notare il modo necessario. Tale valore non tende mai a zero. Supponiamo $\alpha < 9/8$ e studiamo la differenziabilità. Non conviene sare il teorema del differenziale totale quindi calcoliamo le derivate parziali. Chiaramente

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

Mentre

$$\begin{aligned}
\frac{\partial f}{\partial x}(0,0) &= \lim_{x\to 0} \frac{f(x,0) - f(0,0)}{x} \\
&= \lim_{x\to 0} \frac{|x|^{1/4}\ln(1+x^2)}{x|x|^{2\alpha}} \\
&= \lim_{x\to 0} \operatorname{sgn} x \cdot |x|^{\frac{1}{4}+1-2\alpha} \\
&= \lim_{x\to 0} \operatorname{sgn} x \cdot |x|^{\frac{5}{4}-2\alpha}
\end{aligned}$$

che tende a zero per $\alpha < \frac{5}{8}$, altrimeti il limite non esiste. Quindi se $\alpha \geq \frac{5}{8}$ la derivata parziale in 0 non esiste. Sia $\alpha < 5/8$, dove abbiamo il gradiente. Calcoliamo quindi questo limite.

$$\begin{aligned}
\frac{f(h,k) - f(0,0) - \nabla f(0,0)\cdot(h,k)}{\sqrt{h^2+k^2}} &= \frac{f(h,k)}{\sqrt{h^2+k^2}} \\
&= \frac{|x|^{1/4}\ln(1+h^2+k^2)}{(h^2+k^2)^{\alpha+1/2}}
\end{aligned}$$

che tende a zero se e solo se $\alpha + \frac{1}{2} < 9/8$ come detto prima, quindi $\alpha < 5/8$.

Esercizio

Studiare se la seguente funzione è differenziabile nell'origine

$$f(x,y) = \begin{cases} e^{-\frac{1}{x^2}} \sin\sqrt{x^2+y^2} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

Teorema Teorema di Lagrange

Sia $\Omega \subseteq \mathbb{R}^n$ open and $[a,b] \subseteq \Omega$. Let $f\colon \Omega \to \mathbb{R}$ differentiable in Ω. Then, $\exists \alpha \in (a,b)$ such that

$$f(b) - f(a) = \nabla f(\alpha)(b-a)$$

Proof

Costruiamo una funzione che parametrizza il segmento $\varphi\colon [0,1] \to \Omega$ e quindi

$$\varphi(t) = a + t(b-a)$$

che è differenziabile. COnsideriamo la funzione $F = f\circ\varphi\colon [0,1] \to \mathbb{R}$. F è continua in $[0,1]$ ed è differenziabile in $(0,1)$. Ci ricolleghiamo quindi al teorema di Lagrange in una variabile.

1. $F(1) - F(0) = F'(\theta)(1-0)$;
2. $F(1) = f(\varphi(1)) = f(b)$;
3. $F(0) = f(a)$;
4. $F'(t) = (f\circ\varphi)'(t) = Jf(\varphi(t))J\varphi(t) = \nabla f(\varphi(t))\varphi'(t)$;
5. $F'(\theta) = \nabla f(\varphi(\theta))\varphi'(\theta) = \nabla f(\varphi(\theta))(b-a)$, basta porre $a = \varphi(\theta)$.

È falso per le funzioni a valori vettoriali.

Teorema Teorema dell'incremento finto

Sia $\Omega \subseteq \mathbb{R}^n$ aperto e sia $f\colon \Omega \to \mathbb{R}^m$ differenziabile in Ω. Sia $[x, y] \subseteq \Omega$. Allora

$$||f(x) - f(y)|| \leq ||x - y|| \sup_{0<t<1} ||df(x + t(y - x))||$$

Proof Teorema dell'incremento finito

Consideriamo $n = 1$. Il teorema dice che data $g\colon [a, b] \to \mathbb{R}^m$ continua su $[a, b]$ e differenziabile su (a, b). Se $||g'(x)|| \leq \alpha$ per tutte le x in (a, b), allora

$$||g(b) - g(a)|| \leq \alpha(b - a)$$

Poniamo $v = g(b) - g(a)$ e definiamo la funzione ausiliaria $\varphi(x) = g(x) \cdot v$ che ha forma $[a, b] \to \mathbb{R}$. φ è continua su $[a, b]$ e differenziabile su (a, b). Per il tereoam di Lagrange $\exists \theta \in (a, b)$ tale che $\varphi(b) - \varphi(a) = \varphi'(\theta)(b - a)$. Quindi

$$g(b) \cdot v - g(a) \cdot v = g'(\theta) \cdot v(b - a) \leq ||g'(\theta)|| \cdot ||v|| \cdot (b - a)$$

per Cauchy-Schwarz. Usimao il fatto che $v = g(b) - g(a)$. Sostituendo otteniamo

$$||g(b) - g(a)||^2 \leq ||g'(\theta)|| \cdot ||g(b) - g(a)||(b - a)$$

Se $g(b) = g(a)$ la disuguaglianza è automaticamente verificata. Supponiamo $g(b) \neq g(a)$, dividendo per $||g(b) - g(a)||$ ottiene la tesi. Consideriamo ora il naso $n > 1$. Consideriamo la funzione ausiliaria $\Psi(t) = f(x + t(y - x))$ che ha forma $[0, 1] \to \mathbb{R}^m$ continua in $[0, 1]$ e differenziabile in $(0, 1)$. Per ciò che abbiamo appena mostrato sappiamo che $||\Psi(1) - \Psi(0)|| \leq \alpha(1 - 0)$ for

$$\alpha = \sup_{0<t<1} ||\Psi'(t)||$$

Calcoliamo

$$\Psi'(t) = df(x + t(y + x))(y - x)$$

quindi otteniamo

$$||f(y) - f(x)|| \leq \sup_{0<t<1} ||df(x + t(y - x))(y - x)|| \leq ||y - x|| \leq ||y - x|| \sup_{0<t<1} ||df(x + t(y - x))||$$

Corollario

Let $f\colon \Omega \to \mathbb{R}^m$, $\Omega \subseteq \mathbb{R}^n$ aperto connesso, f differenziabile in Ω. Allora f è costante in Ω se e solo se

$$\frac{\partial f_i}{\partial x_j} = 0$$

in Ω per tutte le i, j.

Proof

$(\Longrightarrow)$ Ovvia

($\Longleftarrow$) Sia $x_0 \in \Omega$ e poniamo $E_{x_0} = \{x \in \Omega \,|\, f(x) = f(x_0)\}$. Vogliamo mostrare che $E_{x_0} = \Omega$. Basta mostrare che $E_{x_0} \neq \emptyset$ in quanto $x_0 \in E_{x_0}$ e che E_{x_0} è aperto e chiuso in $\mathbb{R}^m$. Mostriamo che E_{x_0} è aperto. Sia $y \in E_{x_0}$. Sicuramente $\exists r > 0$ tale che $B_r(y) \subseteq \Omega$. Sia $x \in B_r(y)$. $B_r(y)$ è convessa e quindi $[x, y] \subseteq B_r(v)$. Per il torema dell'incremento finito

$$||f(x) - f(y)|| \leq ||x - y|| \sup_{0<t<1} ||df(x + t(y - x))|| = 0$$

perché $df = 0$. Quindi $f(x) = f(y) = f(x_0)$ e $x \in E_{x_0}$, cioè y è interno e quindi E_{x_0} è aperto.

3.3 Criteri di lipschitzianità

Sappiamo che per le funzioni di una variabile la derivata limitata è un criterio di lipschitzianità.

Sia $\Omega \subseteq \mathbb{R}^m$ e $f\colon \Omega \to \mathbb{R}^m$ differenziabile.

Proposition

Sia $f \in \mathrm{Lip}_\alpha(\Omega, \mathbb{R}^m)$. Allora

$$||df(x)|| \leq \alpha$$

per tutte le x in Ω.

Proof

Dobbiamo stimare la norma operatoriale del differenziabile. Basta mostrare che

$$||df(x)v|| \leq \alpha ||v||$$

per tutti i v. Basta farlo per ogni versore $||v|| = 1$. Il differenziale è dato dalla derivata direzionale

$$\begin{aligned} ||df(x)v|| = ||D_v f(x)|| &= \left|\left| \lim_{t\to 0} \frac{f(x+tv) - f(x)}{t} \right|\right| \\ &= \lim_{t\to 0} \left|\left| \frac{f(x+tv) - f(x)}{t} \right|\right| \\ &= \lim_{t\to 0} \frac{1}{|t|} ||f(x+vt) - f(x)|| \leq \alpha \lim_{t\to 0} \frac{1}{|t|} ||tv|| = \alpha ||v|| \end{aligned}$$

che esiste per ipotesi. Siccome la norma è continua possiamo portarla dentro al limite

Aggiungendo un ipotesi vale anche il viceversa in quanto dobbiamo usare il teorema dell'increment finito. Dobbiamo richiedere che tutti i segmenti siano in Ω, quindi ci serve Ω convesso.

Proposition

Sia $\Omega \mathbb{R}^n$ aperto, $f\colon \omega \to \mathbb{R}^m$ e sia $C \subseteq \Omega$ convesso. Supponiamo che $\exists \alpha > 0$ tale che

$$||df(x)|| \leq \alpha, \quad \forall x \in C$$

Allora, $f \in \mathrm{Lip}_\alpha(C)$.

Proof

Abbiamo

$$||f(x) - f(y)|| \leq ||x - y|| \sup_{0<t<1} ||df(x + t(y - x))||$$

Abbiamo una cosa simile alla lipschitzianità ma non è una costante di Lipschitz in quanto dipende da x e y. Però per ipotesi

$$\sup_{0<t<1} ||dt(x+t(y-x))|| \leq \alpha$$

quindi è lipschitziana

Esempio

Sia $f(t) = (\cos t, \sin t, \arctan t)$ da $\mathbb{R}$ a $\mathbb{R}^3$. Verifichiamo se è lipschitziana e le migliori costanti. Nel dominio consideriamo la norma classica mentre nel codominio consideriamo le norme $1, 2, \infty$. Calcoliamo la matrice

$$Jf(t) = f'(t) = \begin{pmatrix} -\sin t \\ \cos t \\ \frac{1}{1+t^2} \end{pmatrix}$$

Dobbiamo calcolare la norma $||Jf(t)||$. Con la norma due abbiamo abbiamo

$$||f(t)||_2 = \sqrt{1 + \left(\frac{1}{1+t^2}\right)^2}$$

Il massimo è chiramente $\sqrt{2}$ che è la miglior costante di Lipschitz. Con la norma 1 abbiamo abbiamo

$$||f(t)||_1 = |\sin t| + |\cos t| + \frac{1}{1+t^2}$$

Il valore del sup è difficile da trovare, possiamo stimarla ad essere minore di 3. Inoltre, $||f'(0)||_1 = 2$ quindi possiamo ridurre il bound. Con la norma infinito abbiamo

$$||f(t)||_\infty = \max\{|\sin t|, |\cos t|, \frac{1}{1+t^2}\}$$

Ci serve il supremum su $t \in \mathbb{R}$ del massimo di queste 3 funzioni. Sicuramente è minore o uguale di 1 in quanto le 3 funzioni lo sono. In $t = 0$ otteniamo precisamente 1. Quindi il supremum è precisamente 1. È di lipschitz (sempre) ma le costanti dipendono dalla norma, e non è mai una contrazione.

4 Derivate di ordini successive

Prendiamo $\Omega \subseteq \mathbb{R}^n$ aperto e consideriamo $f\colon \Omega \to \mathbb{R}$. Fissiamo un versore v e prendiamo $a \in \Omega$. Supponiamo che la derivata direzionale $D_v f$ esista in un intorno $B_r(a)$. Allora $D_v f$ è una funzione definita in un aperto che contiene a. Quindi, possiamo farne a sua volta la derivata direzionale. Fissiamo un altro versore $w \in \mathbb{R}^m$ e costruiamo la derivata direzionale nella direzione di w nel punto a se esiste.

$$D_w(D_v f)(a) \triangleq D^2_{w,v} f(a)$$

In particolare, se come direzioni w e v scegliamo le direzioni degli assi cartesiani, queste derivate direzionali si chiamano derivate parziali seconde. Essa viene denotata come

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) \triangleq \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}(a)\right) \triangleq f_{x_j x_i}(a)$$

Se $i \neq j$ si parla di derivate parziali secondo miste. Se $i = j$ si usa la notazione

$$\frac{\partial^2 f}{\partial x_i^2}(a)$$

Possiamo quindi costruire n^2 derivate parziali, con cui costruiamo una matrice quadrata.

Definizione Matrice Hessiana

Quindi, se in a esistono tutte le derivate parziali seconde di f possiamo definire la matrice Hessiana.

$$Hf(a) \triangleq \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(a) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(a) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(a) \\ \vdots & & & \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(a) & \frac{\partial^2 f}{\partial x_n \partial x_2}(a) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(a) \end{pmatrix}$$

Esempio

Sia $f(x, y, z) = 3x^2 y + xz^3$. Allora

$$Hf(x, y, z) = \begin{pmatrix} 6y & 6x & 3z^2 \\ 6y & 0 & 0 \\ 3z^2 & 0 & 6xz \end{pmatrix}$$

Esempio

Sia $f(x, y) = x^2 \sqrt[3]{y}$. Abbiamo problemi di differenziabilità nell'asse x. Ma per $y \neq 0$ vale

$$\nabla f = \left(2x\sqrt[3]{y}, \frac{1}{3}\frac{x^2}{\sqrt[3]{y}}\right)$$

Fuori dall'asse x la funzione è differenziabile per il teorema del differenziale totale. Consideriamo un punto generico in $(x_0, 0)$ dove f è identicamente nulla. Quindi

$$\frac{\partial f}{\partial x}(x_0, 0) = 0$$

L'altra invece dobbiamo calcolare

$$\begin{aligned} \frac{\partial f}{\partial y}(x_0, 0) &= \lim_{y \to 0} \frac{f(x_0, y) - f(x_0, 0)}{y} \\ &= \lim_{y \to 0} \frac{x_0^2}{y^{2/3}} \\ &= \begin{cases} \infty & x_0 = 0 \\ 0 & x_0 \neq 0 \end{cases} \end{aligned}$$

Quindi se $x_0 = 0$

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

perché $f(0,y) = 0$. f è differenziabile nell'origine (ma non possoamo mostrare con il teorema del differenziale totale). Dobbiamo calcolare

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

che si risolve in coordinate polari. Non possiamo calcoalre $f_{xy}(0)$ perché dovrei calcolare

$$\lim_{x\to 0} \frac{f_y(x,0) - f_y(0,0)}{x}$$

che none siste. Invece $f_{yx}(0)$ abbiamo

$$\lim_{y\to 0} \frac{f_x(0,y) - f_x(0,0)}{y} = 0$$

Quindi l'esistenza di una derivata parziale mista non implica l'esistenza di quella con gli ordini scambiati.

Esempio

Sia

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

Abbimamo che f è continua in $0,0$ in quanto in coordinate polari

$$\left| \frac{\rho\cos\theta\rho^2\sin^3\theta}{\rho^2} \right| \leq \rho^2 \to 0$$

Calcoliamo il gradiente di f fuori dall'origine (ci serve per calcolare le derivate secondarie). Per $(x,y) \neq (0,0)$ abbiamo

$$f_x(x,y) = \frac{y^3(x^2+y^2) - xy^3 \cdot 2x}{(x^2+y^2)^2} = \frac{y^3(y^2-x^2)}{(x^2+y^2)^2}$$

e

$$f_y(x,y) = \frac{3xy^2(x^2+y^2) - xy^3 \cdot 2y}{(x^2+y^2)^2} = \frac{xy^2(3x^2+y^2)}{(x^2+y^2)^2}$$

f è sicuramente differenziabile duroi dall'origine. Inoltre $\nabla f(0,0) = (0,0)$ perché f è costante negli assi. Per il teorema del differenziale totale f è differenziabile in $\mathbb{R}^2$. Calcoliamo

$$f_{xy}(0,0) = \lim_{x\to 0} \frac{f_y(x,0) - f_y(0,0)}{x} = 0$$

perché $f_y(x,0) = 0$. Invece,

$$\begin{aligned} f_{yx}(0,0) &= \lim_{y\to 0} \frac{f_x(0,y) - f_x(0,0)}{x} \\ &= \lim_{y\to 0} \frac{y^5/y^4 - 0}{y} \\ &= \lim_{y\to 0} 1 = 1 \end{aligned}$$

Quindi anche se ambo le derivate miste esistono non coincidono necessariamente.

Teorema Teorema di Schwarz

Sia $f\colon \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$, Ω aperto, $a \in \Omega$ e v, w due versori in $\mathbb{R}^m$. Supponiamo che $B_r(a)$ dove $D^2_{v,w}f$ e $D^2_{w,v}f$ esistono e siano continue in a. Allora $D^2_{v,w}f(a) = D^2_{w,v}f(a)$.

Corollario

Se le derivate parziali esistono e sono continue in Ω, allora

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$$

ovvero $Hf(x)$ è simmetrica.

Definizione Classe $\mathcal{C}^2$

Sia $f\colon \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$, allora $f \in \mathcal{C}^2$ se tutte le derivate parziali prime e seconde di f esistono e sono continue in Ω.

Sia $f\colon \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^m$. Possiamo definire le derivate superiori analogamente. Se f è differenziabile in Ω è definito il differenziale $df(x)\colon \mathbb{R}^n \to \mathbb{R}^m$ dato da $df(x)h = Jf(x)h$. Fissato x, al variare di h ho un applicazione lineare

$$df\colon \Omega \to \mathcal{L}(\mathcal{R}^n, \mathcal{R}^m)$$

che manda $x \to df(x)$. Questa non è in generale lineare in x.

Definizione

Sia $f\colon \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^m$ è differenziabile due volte in $a \in \Omega$ se f è differenziabile in Ω e $df\colon \Omega \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ è differenziabile in a.

Proposition

1. f è differenziabile due volte in a se e solo se $\partial f / \partial x_i$ è differenziabile in a per tutti gli i.
2. se f è differenziabile due volte in a, allora esistono tutte le seconde derivate parziali e posso scambiare l'ordine di derivazione.
3. se f è differenziabile due volte in a, allora
   (a) $D_v f$ è differenziabile in a per ogni v;
   (b) $D^2_{v,w}f(a) = D^2_{w,v}f(a)$;
   (c)
   $$D^2_{v,w}f(a) = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(x) v_i w_i$$
4. Se in un intorno di a esistono e sono continue tutte le derivate parziali, esistono tutte le derivate parziali seconde e sono continue in a, allora f è due volte differenziabile in a.
5. se $f \in \mathcal{C}^2(\Omega)$ allora f è due volte differenziabile in Ω.

Definizione

Se $f\colon \Omega \to \mathbb{R}^m$ è due volte differenziabile in a, allora $d^2 f(a)$ indica il differenziale del differenziale di f nel punto a.

Quindi

$$d^2 f(a) \in \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m))$$

Questo significa che $d^2f(a)h \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$. Cioè è lineare sia in h che in k, quindi bilineare $k \to (d^2f(a)h)k$. Siccome è bilinare possiamo vederlo come $d^2f(a)\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$ dato da $d^2f(a)(h,k) = (d^2f(a)h)k$. Possiamo dire che

$$d^2f(a)(h,k) = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a)h_i k_j = h^t Hf(a)k = (Hf(a)k)h$$

dove $Hf(a)$ è un tensore. In particolare, se f è una funzione a valori reali

$$df(a)h = \nabla f(a)h, \quad d^2f(a)(h,k) = h^t Hf(a)k$$

è una forma bilineare (forma = valori reali). Quando $h = k$ scriviamo una sola variabile. In tal caso ottengo una forma quadratica che si scrive

$$\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a)h_i j_j$$

Succede che per $h \to 0$

$$\frac{d^2f(a)(h)}{||h||^2}$$

è limitato.

Esempio

Data una forma quadratica $ax^2 + bxy + cy^2$ e consideriamo

$$\frac{ax^2 + bxy + xy^2}{ax^2 + bxy + xy^2}$$

che in coordinate polari diventa

$$a\cos^2\theta + b\sin\theta\cos\theta + c\sin^2\theta$$

che chiaramente è limitata da $|a| + |b| + |c|$.

4.1 Derivate k-esime

Se per $\{v^{(i)}\}$ versori in $\mathbb{R}^m$ esiste in $B_\delta(a)$ le derivate direzionali

$$D^{k-1}_{v^{(k-1)},\cdots,v^{(1)}} f(x)$$

e se è derivabile in a lungo la direzione $v^{(k)}$, allora diremo che esiste la derivata k-esima. Se esistono continue tutte le derivate k esime possiamo permutare gli indici.

Esempio

Con $k = 3$ sia $f\colon \mathbb{R}^2 \to \mathbb{R}$. Possiamo scrivere $f_{xxx}, f_{xxy}, f_{xyx}, f_{yxx}, f_{yyx}, f_{yxy}, f_{xyy}, f_{yyy}$ che sono 2^3 derivate. Se vale il teorema di Schwarz molte di queste coincidono.

Definizione

Diciamo che $f \in \mathcal{C}^k(\Omega)$ se ogni derivata parziale di ordine $\leq k$ esiste ed è continua in Ω.

Chiaramente sono annidati.

4.2 Differenzaibili k-esimi

Definizione Differenzaibile k-esimi

$$d^k f(a)\colon \mathbb{R}^n \times \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}^n$$

è un applicazione lineare in ciascuno dei vettori di ingresso definita da

$$d^k f(a)(h^{(1)}, \cdots, h^{(k)}) = \sum_{i_1, \cdots, i_k = 1}^{n} \frac{\partial^k f}{\partial x_{i_1} \cdots \partial x_{i_k}}(a) h_{i_i}^{(i)}$$

Per Schwarz molte di queste derivate sono uguali. Ci interessa il caso in cui $h^{(i)} = h^{(j)} = h$ In questo caso abbiamo

$$d^k f(a)(h) = \sum_{\substack{j_1, \cdots, j_n \geq 0 \\ \sum j_i = k}} \binom{k}{j_1 \cdots j_n} \frac{\partial^k f}{\partial x_n^{j_n} \cdots \partial x_1^{j_1}}(a) h_1^{j_1} \cdots h_n^{j_n}$$

dove abbiamo usato il coefficiente multinomiale.

Esempio
Sia $f\colon \mathbb{R}^3 \to \mathbb{R}$ e $k = 4$. Allora

$$d^4 f(a)(h) = f_{xxxx} h_1^4 + f_{yyyy} h_2^4 + f_{zzzz} h_3^4 + h_1^3 h_2 \left(f_{xxxy} + f_{xxyx} + f_{xyxx} + f_{yxxx}\right) + \cdots$$

Il termine nelle parentesi è pari a $4 f_{xxxx}$. E il valore

$$\binom{4}{3, 1, 0} = \frac{4!}{3!1!0!} = 4$$

$d^4 f(a)(h)$ è un polinomio omogeneo di grado k in $h_1, \cdots, h_n$. Inoltre valore che

$$\frac{d^4 f(a)(h)}{||h||^k}$$

è limitato per $h \to 0$.

Teorema Teorema di Taylor
Sia $\Omega \subseteq \mathbb{R}^n$ aperto, siano $a, x \in \Omega$ tale che $[a, x] \subseteq \Omega$. Sia $f\colon \Omega \to \mathbb{R}$ tale che $f \in \mathcal{C}^k(\Omega)$. Allora,

$$f(x) = \sum_{i=0}^{k-1} \left(\frac{d^i f(a)(x-a)}{i!} \right) + R_k(x-a)$$

è il polinomio di Taylor di ordine $k-1$ di f in a. Il resto è dato in forma di Peano come

$$R_k(x-a) = \frac{d^k f(a)(x-a)}{k!} + o(||x-a||^k)$$

per $x \to a$, oppure in forma di Lagrange (se al posto di fermarmi a $k-1$ ci fossimo fermati a k)

$$R_k(x-a) = \frac{d^k f(y)(x-a)}{k!}$$

con $y \in (a, x)$. Inoltre, il polinomio di Taylor è l'unico polinomio di grado al più k tale che $o(||x-a||^k)$ per $x \to a$.

Viene detto ordine del polinomio di Taylor e non grado in quanto il polinomio ha grado al più quello dato ma potrebbe essere minore.

Esempio

Sia
$$f(x, y, z) = \ln(1 + xz^2 - y^3 e^{x+1}) + \sin\sqrt{1 - x^2 + xy - z\ln(1 + xy)}$$
Trovare lo sviluppo di MacLaurin al quarto ordine. Poniamo $t = xz^2 - y^3 e^{x+1}$, e osserviamo che $t \to 0$ se $(x, y, z) \to (0, 0, 0)$. Possiamo dire

$$\begin{aligned} t &= xz^2 - y^3 e\left[1 + x + \frac{x^2}{2} + \cdots\right] \\ &= xz^2 - ey^3 - exy^3 + o(||(x, y, z)||^4) \end{aligned}$$

Sviluppiamo ora il logaritmo

$$\ln(1 + t) = xz^2 - ey^3 - exy^3 + o(||(x, y, z)||^4)$$

in quanto tutti gli altri termini sono trascurabili. Esercizio: fare anche il seno e mettere assieme.

5 Ottimizzazione

Sia $\Omega \subseteq \mathbb{R}^n$ un sottoinsieme qualcunque e sia $f\colon \Omega \to \mathbb{R}$.

Definizione

Sia $a \in \Omega$ è un punto di massimo assoluto per f in Ω se $f(x) \leq f(a)$ $\forall x \in \Omega$. Invece a è punto di massimo relativo (o locale) per f se esiste U intorno di a tale che a è punto di massimo assoluto per f in $\Omega \cap U$. Analogamente forte per il minore stretto.

Se Ω è un aperto, il problema di cercare massimi e minimi di f in Ω si chiama ottimizzazione libera.

Teorema Teorema di Fermat

Sia $\Omega \subseteq \mathbb{R}^n$ aperto, $a \in \Omega$ un punto estremante per f. Se f è differenziabile in a, allora

$$\frac{\partial f}{\partial x_i} = 0$$

Cioè $\nabla f = 0$.

Proof Teorema di Fermat

Supponiamo che a sia estremanete per f. Allora a è estremanete anche a tutti le restrizioni di rette che passano per a. In particolare prendiamo $\gamma_i(t) = a + te_i$ paralleli agli assi cartesiani. Sia Φ_i la restrizione di f a tali rette. Quindi $Phi_i = f \circ \gamma_i$. Le funzioni Φ_i hanno un punto estremante in $t = 0$. Inoltre, Φ_i sono differenziabili in quanto composizione di funzioni differenziabili. Allora, per il teorema di Fermat in una variabile vale $\Phi'(0) = 0$. Quindi abbiamo

$$\begin{aligned}\Phi_i'(0) &= (f \circ \gamma_i)'(0) = \nabla f(\gamma_i(0)) \circ \gamma_i'(0) \\ &= \nabla f(a) \cdot e_i = \frac{\partial f}{\partial x_i}(a) = 0\end{aligned}$$

Tale condizione è necessaria ma non sufficiente.

Esempio

Siano $f(x,y) = x^2 + y^2$ e $g(x,y) = x^2 - y^2$. Per entrambe $\nabla f(0,0) = \nabla g(0,0) = 0$. Tuttavia, l'origine è estremante per f ma non per g. Mentre non lo è per g, che ha un punto di sella.

Definizione Punti stazionari

Se f è differenziabile in $a \in \Omega$ e se $\nabla f(a) = 0$, diremo che a è punto stazionario o punto critico per f.

Definizione Punto di sella

Se f è differenziabile in $a \in \Omega$ e a è un punto stazionario che non è estremante, allora è un *punto di sella*.

Vogliamo avere strumenti per trovare questi punti. Non ha senso studiare il segno della derivata in quanto ce ne sono molteplici. Possiamo considerare i differenziali secondi, che sono forme bilineari, e studiare il segno della forma quadratica associata.

Per le funzioni di una variabile possiamo considerare l'espansione di Taylor e studiare l'incremento (dove il primo termine si annulla siccome siamo in un intorno di un punto stazionario)

$$f(x) - f(a) = (x-a)^2 \left[\frac{1}{2} f''(a) + o(1)\right]$$

da cui seguono i casi del segno. Vogliamo fare un procedimento analogo. Sia f differenziabile 2 volte Allora per $h \to 0$

$$f(a+h) = f(a) + \nabla f(a) \cdot h + \frac{1}{2} h^t Hf(a)h + o(||h||^2)$$

Se a è punto critico, l'incremento di f ha la forma

$$f(a+h) - f(a) = \frac{1}{2} h^t Hf(a)h + o(||h||^2) = ||h||^2 \left[\frac{h^t Hf(a)h}{||h||^2} + o(1) \right]$$

Dobbiamo quindi studiare il segno della funzione

$$F(h) = \frac{h^t Hf(a)h}{||h||^2} = \frac{1}{||h||^2} \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a) h_i h_j$$

F è il quoziente di due polinomi omogenei di grado 2. Quindi $F \colon \mathbb{R}^n \backslash \{0\} \to \mathbb{R}$ è omogenea di grado zero $F(\lambda h) = F(h)$. Ciò significa che F è costante su tutti i punti di ciascuna retta passante per l'origine (a parte nell'origine). Vogliamo dimostrare che F ammette massimi e minimi assoluti. Notiamo che se restringiamo il dominio ad un disco (non-degenere), i valori che questa funzione assume rimangono gli stessi, non importa quanto piccolo il cerchio. Ci restringiamo quindi a

$$S = \{x \in \mathbb{R}^n \mid ||x|| = r\}$$

e allora F rispetto a S cioè $S \to \mathbb{R}$ ammette massimo e minimo assoluti, in quanto F è continua e S è compatto, quindi per Weierstrass. Allora esistono $m, M \in \mathbb{R}$ tale che $m \leq F(h) \leq M$ per tutte le $h \in \mathbb{R}^n \backslash \{0\}$ e m, M sono

$$m \leq \frac{h^t Hf(g)h}{||h||^2} \leq M$$

per tutte le $h \in \mathbb{R}^n \backslash \{0\}$ e quindi

$$m||h||^2 \leq h^t Hf(a)h \leq M||h||^2$$

per tutte le $h \in \mathbb{R}^n$. Possiamo quindi dedurre che se $m > 0$ allora a è un punto di minimo relativo, se $M < 0$ allora a è un punto di massimo relativo, se invece $m < 0 \wedge M > 0$ allora $h^t Hf(a)h$ cambia segno e quindi a è un punto di sella. Se $mM = 0$, c'è almeno una retta lungo la quale il termine di secondo grado si annulla. In questo caso $o(1)$ non è più trascurabile e non possiamo concludere niente.

Esempio

Sia

$$f(x,y) = (x-y)^2 + \varphi$$

dove φ è un polinomio composto da monomi di grado maggiore o uguale a tre. Chiaramente f è differenziabile in 0 e $f(0,0) = 0$ quindi è punto critico. Calcoliamo

$$\begin{aligned} f_x &= 2(x-y) + \varphi_x \\ f_y &= -2(x-y) + \varphi_y \\ f_{xx} &= 2 + \varphi_{xx} \\ f_{xy} &= -2 + \varphi_{xy} \\ f_{yy} &= 2 + \varphi_{yy} \end{aligned}$$

Molte di queste derivate nell'origine scompaiono in dato il grado del polinomio. Quindi

$$Hf(0,0) = \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix}$$

Scegliendo $h = (1,1)$ abbiamo che

$$\begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 0$$

Scegliendo $\varphi(x,y) = x^{16}$. Con questa scelta f ha un minimo in $0,0$. Se però scelgo $\varphi(x,y) = x^{15}$ questa ha un punto di sella nell'origine.

Studiamo $h^t Hf(a)h$. In generale, se $Q\colon \mathbb{R}^n \to \mathbb{R}$ è una forma quadratica, allora $Q(v) = v^t Av$ dove A è una matrice $n \times n$ simmetrica. Per il teorema spettrale reale, se A è simmetrica, allora A ha n autovalori reali

$$\lambda_1 \leq \cdots \leq \lambda_n$$

ed esiste una rotazione R tale che $A = R^t DT$ dove $D = \text{dial}(\lambda_1, \cdots, \lambda)$. Quindi,

$$\begin{aligned} Q(v) &= v^t Av = v^t R^t DRv \\ &= \tilde{v}^t D \tilde{v} \\ &= \sum \lambda_i \tilde{v}_i^2 \end{aligned}$$

ponento $\tilde{v} = Rv$. Quindi

$$\lambda_1 ||\tilde{v}||^2 \leq Q(v) \leq \lambda_n ||\tilde{v}||^2$$

Siccome R è una rotazione, quindi un isometria, preserva la norma, quindi $||\tilde{v}|| = ||Rv|| = ||v||$. Allora

$$\lambda_1 ||v||^2 \leq Q(v) \leq \lambda_n ||v||^2$$

Possiamo quindi dedurre che se $\lambda_1 > 0$, Q è definita positiva. Se $\lambda_m < 0$, Q è definita negativa. Se $\lambda_1 \lambda_n < 0$, allora Q è indefinita, se $\lambda_1 \lambda_n = 0$, allora Q è semidefinita.

Questa era la dimostrazione del seguente teorema.

Teorema

Sia $f\colon \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$, Ω aperto e $a \in \Omega$ punto critico per f ed f differenziabile due volte in a. Allora,

1. $Hf(a)$ è definita positiva, allora a è punto di minimo relativo forte;
2. $Hf(a)$ è definita negativa, allora a è punto di massimo relativo forte;
3. $Hf(a)$ è indefinita, allora a è un punto di sella;
4. se a è punto di massimo relativo, allora $Hf(a)$ è semidefinita negativa;
5. se a è punto di minimo relativo, allora $Hf(a)$ è semidefinita positiva.

Se la matrice Hessiana è semidefinita, non possiamo fare nulla.

Per $n = 2$ possiamo determinare il segno di $Hf(a)$ senza trovare gli autovalori

$$\begin{pmatrix} h & k \end{pmatrix} \begin{pmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{pmatrix} \begin{pmatrix} h \\ k \end{pmatrix} = h^2 f_{xx} + 2hk f_{xy} + k^2 f_{yy}$$

Se $f_{xx} = f_{yy} = 0$ allora la forma quadratica è indefinita. Supponiamo quindi che $f_{xx} \neq 0$. Abbiamo quindi

$$\begin{aligned} h^2 f_{xx} + 2hk f_{xy} + k^2 f_{yy} &= f_{xx} \left(h^2 + 2hk \frac{f_{xy}}{f_{xx}} \right) + k^2 f_{yy} \\ &= f_{xx} \left(h^2 + 2hk \frac{f_{xy}}{f_{xx}} + k^2 \frac{f_{xy}^2}{f_{xx}^2} \right) - k^2 \frac{f_{xy}^2}{f_{xx}} + k^2 f_{yy} \\ &= f_{xx} \left(h + k \frac{f_{xy}}{f_{xx}} \right)^2 + \frac{k^2}{f_{xx}} \left(f_{xx} f_{yy} - f_{xy}^2 \right) \\ &= f_{xx} \left(h + k \frac{f_{xy}}{f_{xx}} \right)^2 + \frac{k^2}{f_{xx}} \det Hf(a) \end{aligned}$$

il segno di quest'ultimo è completamente determinato dal segno del determinante e di f_{xx}.

Teorema

Sia $\Omega \subseteq \mathbb{R}^2$ aperto, $f\colon \Omega \to \mathbb{R}$ due volte differenziabile, a punto critico. Allora:

1. se $\det Hf(a) < 0$ allora a è punto di sella;
2. se $\det Hf(a) > 0$ e $f_{xx} > 0$ allora a è punto di minimo;
3. se $\det Hf(a) > 0$ e $f_{xx} < 0$ allora a è punto di massimo;

Se è nullo il criterio è inutile.

Notiamo che se il determinante è positivo allora f_{xx} e f_{yy} hanno lo stesso segno. Infatti $f_{xx}f_{yy} - f_{xy}^2 > 0$ e quindi $f_{xx}f_{yy} > f_{xy}^2 \geq 0$.

Per $n = 2$ il segno di $Hf(a)$ può essere determinato mediante il criterio dei minori principali di nord-ovest.

Esempio

Sia $f(x, y) = x^2 - 2y^2 - 4x + 7$. Chiaramente $f \in \mathcal{C}^\infty(\mathbb{R}^2)$. Abbiamo $f_x = 2x - 4$ e $f_y = -4y$, $f_{xx} = 2$ e $f_{yy} = -4$. Le prime due derivate si annullano per

$$\begin{cases} 2x - 4 = 0 \\ -4y = 0 \end{cases}$$

quindi in $(2, 0)$. Abbiamo

$$Hf(a)(2, 0) = \begin{pmatrix} 2 & 0 \\ 0 & -4 \end{pmatrix}$$

quindi $(2, 0)$ è punto di sella.

Esempio

Sia $f(x, y) = x^4 + y^4 + 4xy - 2x^2 - 2y^2$. Chiaramente $f \in \mathcal{C}^\infty(\mathbb{R}^2)$. Abbiamo $f_x = 4x^3 + 4y - 4x$ e $f_y = 4y^3 + 4x - 4y$.

$$\begin{cases} x^3 + y - x = 0 \\ y^3 + x - y = 0 \end{cases} \to \begin{cases} x^3 + y^3 = 0 \\ y^3 + x - y = 0 \end{cases} \to \begin{cases} x = -y \\ y^3 - 2y = 0 \end{cases}$$

quindi i punti sono $(0, 0), (\sqrt{2}, -\sqrt{2})$ e $(-\sqrt{2}, \sqrt{2})$. Abbiamo $f_{xx} = 12x^2 - 4$ e $f_{xy} = 4$ e $f_{yy} = 12y^2 - 4$. Dato

$$Hf(x, y) = \begin{pmatrix} 12x^2 - 4 & 4 \\ 4 & 12y^2 - 4 \end{pmatrix}$$

Per i punti con le radici

$$Hf(\pm\sqrt{2}, \mp\sqrt{2}) = \begin{pmatrix} 20 & 4 \\ 4 & 20 \end{pmatrix}$$

il cui determinante è non-negativo, e pure f_{xx}, quindi questi punti sono di minimo. Nell'origine abbiamo

$$Hf(0, 0) = \begin{pmatrix} -4 & 4 \\ 4 & -4 \end{pmatrix}$$

che ha determinante nullo. In questo caso studiamo l'incremento $f(x, y) - f(0, 0) = f(x, y) = x^4 + y^2 - 2(x - y)^2$. Consideriamo $f(x, -x) = 2x^4 - 8x^2$ che ha un massimo per $x = 0$. COnsideriamo $f(x, x) = 2x^4$ che ha pure un minimo in $x = 0$. Quindi $(0, 0)$ è un punto di sella.

Esempio

Sia $f(x, y) = x^2y - y^2$. Chiaramente $f \in \mathcal{C}^\infty(\mathbb{R}^2)$. Abbiamo $f_x = 2xy$ e $f_y = x^2 - 2y$. I punti critici sono solo l'origine. Abbiamo $f_{xx} = 2y$ e $f_{xy} = 2x$ e $f_{yy} = -2$. Ma matrice è data da

$$Hf(x,y) = \begin{pmatrix} 2y & 2x \\ 2x & -2 \end{pmatrix}$$

e nell'origine

$$Hf(0,0) = \begin{pmatrix} 0 & 0 \\ 0 & -2 \end{pmatrix}$$

ha determinante nullo. Quindi il criterio è inutile. Consideriamo la restrizione di f alle rette passanti per l'origine. $f(x, mx) = mx^3 - m^2x^2$ che ha un punto di massimo per $x = 0$ per tutte le m. Notiamo $f(0, y) = -y^2$ che ha un punto di max in $y = 0$. Ma non è legittimo concludere che f ha un massimo nell'origine. Studiamo il segno dell'incremento $f(x,y) - f(0,0) = f(x,y) = y(x^2 - y)$. Questa funzione è non-negativa se e solo se $y > 0$ e $x^2 > y$ oppure se $y < 0$ e $x^2 < y$. Ma la seconda non è possibile. Quindi f ha un punto di sella nell'origine. Infatti ogni intorno dell'oigine contiene punti in cui $f(x,y) > 0$ e punti in cui $f(x,y) < 0$. Per esempio $f(x, x^2/2)$ ha un minimo per $x = 0$.

La ricerca degli estremanti può essere fatta anche in coordinate polari. Se $F(\rho,\theta) = f(\rho\cos\theta, \rho\sin\theta)$ i punti critici corrispondono a

$$\frac{\partial F}{\partial \rho} = 0 \quad \frac{\partial F}{\partial \theta} = 0$$

per $\rho > 0$. Dobbiamo tuttavia trattare $\rho = 0$ cioè l'origine separatamente.

Esempio

Sia

$$f(x,y) = \frac{x^2 + y^2}{x} - 4\ln(1+y)$$

quindi

$$f_x = 1 - \frac{y^2}{x^2} \quad f_y = \frac{2y + 2y^2 - 4x}{x(1+y)}$$

Si annullano entrambi in $(0,0), (1,1), (3,-3)$ ma l'unico punto accettabile è $(1,1)$. Abbiamo $f_{xx} = 2y^2/x^2$, $f_{xy} = -2y/x^2$ e $f_{yy} = 2/x + 4/(1+y)^2$. Da questo

$$Hf(x,y) = \begin{pmatrix} 2\frac{y^2}{x^3} & -2\frac{y}{x^2} \\ -2\frac{y}{x^2} & \frac{2}{x} + \frac{4}{(1+y)^2} \end{pmatrix}$$

e nel punto

$$Hf(1,1) = \begin{pmatrix} 2 & -2 \\ -2 & 3 \end{pmatrix}$$

il cui determinante è maggiore di zero e $f_{xx} > 0$. Quindi $(1,1)$ è un punto di minimo

6 Teorema della funzione implicita

Sia $\Omega \subseteq \mathbb{R}^2$ aperto e $F\colon \Omega \to \mathbb{R}$ e consideriamo $S = \{(x,y) \mid F(x,y) = 0\}$.