

ANALISI II

Federico Mainetti Gambera

15 dicembre 2019

Calcolo differenziale per funzioni reali di più variabili

Continuità di una funzione in più variabili:

Una funzione $f : \mathbb{R}^n \rightarrow \mathbb{R}$ è continua in un punto x_0 se

$$\lim_{x \rightarrow x_0} f(x) = f(x_0)$$

La continuità di una funzione è anche deducibile dal fatto che sia costituita (somma/ prodotto/ quoziente/ certe volte anche composizione) da funzioni elementari continue.

Calcolo di limiti in più variabili

Non esistenza del limite

Per mostrare che una certa funzione in più variabili non ammette limite in un determinato punto, è sufficiente determinare due curve passanti per il punto lungo le quali la funzione assume limiti diversi.
es.

$$\lim_{(x,y) \rightarrow (0,0)} \frac{xy}{x^2 + y^2}$$

Analizziamo la funzione lungo due curve:

- con $y = x$ ottengo $f(x, x) = \frac{1}{2}$
- con $y = -x$ ottengo $f(x, -x) = -\frac{1}{2}$

non ammette limite.

Uso di maggiorazioni con funzioni radiali per provare l'esistenza del limite

Per dimostrare l'esistenza di un limite per $(x, y) \rightarrow (0, 0)$, si impone $x = \rho \cdot \cos(\theta)$ e $y = \rho \cdot \sin(\theta)$, successivamente si l'intera funzione sotto modulo e procede con semplificazioni e maggiorazioni (per eliminare i seni e i coseni). E' essenziale che la funzione non dipenda da θ .

Più in generale se si volesse calcolare il limite per $(x, y) \rightarrow (x_0, y_0)$ si pongono $x = x_0 + \rho \cdot \cos(\theta)$ e $y = y_0 + \rho \cdot \sin(\theta)$

Note sugli esercizi

- Se il limite non presenta una forma di indeterminazione allora il valore cercato si ricava sostituendo direttamente il punto nella funzione.
- Tecniche standard della maggiorazione:

– disuguaglianza triangolare:

$$|a + b| \leq |a| + |b|$$

– maggiorazione di frazioni, con $a, b, c \geq 0$:

$$\frac{a}{b+c} \leq \frac{a}{b}$$

– maggiorazione di funzioni trigonometriche:

$$|\cos(\theta)| \leq 1, \quad |\sin(\theta)| \leq 1$$

- Il criterio che ci permette di trovare il limite richiede di trovare una funzione maggiorante di $|f|$ che sia radiale (dipenda solo da ρ , non θ) e infinitesima. Da notare è che è anche possibile semplificare la funzione anche senza passare subito in coordinate polari.

- Solitamente si suddivide la funzione in una serie di somme di funzioni e si studiano quest'ultime separatamente.

- Spesso la formula

$$\cos^2(\theta) + \sin^2(\theta) = 1$$

risulta molto utile.

Topologia in \mathbb{R}^n e proprietà delle funzioni continue

Sia E un sottoinsieme di \mathbb{R}^n , un punto x_0 si dice:

- interno ad E , se esiste un intorno centrato in x_0 contenuto in E ;
- esterno ad E , se esiste un intorno centrato in x_0 contenuto in E^c ;
- di frontiera per E , se ogni intorno centrato in x_0 contiene almeno un punto di E e uno di E^c .

Un insieme $E \subseteq \mathbb{R}^n$ si dice:

- aperto, se ogni suo punto è interno a E ;
- chiuso, se il suo complementare è aperto.

Sia E un sottoinsieme di \mathbb{R}^n , si dice:

- interno di E , e si indica con E° , l'insieme dei punti interni di E ;
- frontiera o bordo di E , e si indica con δE , l'insieme dei punti di frontiera di E ;
- chiusura di E , e si indica con \bar{E} , l'insieme $E \cup \delta E$.

Alcune informazioni extra:

- si ha sempre $E^\circ \subseteq \delta E \subseteq \bar{E}$;
- il complementare di un aperto è chiuso e viceversa;
- esistono insiemi nè aperti nè chiusi, gli unici insiemi sia aperti sia chiusi sono quello vuoto e \mathbb{R}^n ;
- l'unione di una famiglia qualsiasi (anche infinita) di insiemi aperti e l'intersezione di un numero finito di insiemi aperti sono insiemi aperti
- l'intersezione di una famiglia qualsiasi (anche infinita) di insiemi chiusi è l'unione di un numero finito di insiemi chiusi sono insiemi chiusi;
- un insieme aperto non contiene nessuno dei suoi punti di frontiera, un insieme chiuso contiene tutti i suoi punti di frontiera.

Un insieme si dice:

- limitato se esiste un intorno che lo contiene tutto;
- connesso se per ogni coppia di punti dell'insieme, esiste un arco continuo che li connette contenuto nell'insieme.

Proprietà topologiche delle funzioni continue:

- **teor.** Teorema di Weierstrass. Sia $E \subset \mathbb{R}^n$ un insieme chiuso e limitato e $f : E \rightarrow \mathbb{R}$ sia continua, allora f ammette massimo e minimo in E , cioè esistono x_m e x_M tali che $f(x_m) \leq f(x) \leq f(x_M)$ per ogni $x \in E$.
- **teor.** Teorema degli zeri. Sia E un insieme connesso di \mathbb{R}^n e $f : E \rightarrow \mathbb{R}$ sia continua. Se x, y sono due punti di E tali che $f(x) < 0$ e $f(y) > 0$, allora esiste un terzo punto $z \in E$ in cui f si annulla. In particolare, lungo ogni arco di curva continua contenuto in E che congiunge x e y , c'è almeno un punto in cui f si annulla.

Derivate parziali, piano tangente, differenziale

Derivata parziale

Calcolo di una derivata parziale tramite la definizione di rapporto incrementale in un punto (x_0, y_0) . Per prima fissiamo $y = y_0$ e deriviamo rispetto alla x :

$$\frac{\delta f}{\delta x}(x_0, y_0) = \lim_{h \rightarrow 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h}.$$

Successivamente facciamo l'opposto, cioè fissiamo $x = x_0$ e deriviamo rispetto alla y :

$$\frac{\delta f}{\delta y}(x_0, y_0) = \lim_{k \rightarrow 0} \frac{f(x_0, y_0 + k) - f(x_0, y_0)}{k}.$$

Una funzione $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ si dice derivabile in un punto del suo dominio se in quel punto esistono tutte le sue derivate parziali; si dice derivabile in A se è derivabile in ogni punto di A .

Se f è derivabile in un punto, chiameremo gradiente $(\nabla f(x))$ il vettore delle sue derivate parziali.

Piano tangente

Costruire il piano tangente a una funzione in due variabili in un punto (x_0, y_0) :

1. troviamo la retta tangente alla funzione nel piano $y = y_0$:

$$\begin{cases} z = f(x_0, y_0) + \frac{\delta f}{\delta x}(x_0, y_0) \cdot (x - x_0) \\ y = y_0 \end{cases}$$

2. troviamo la retta tangente alla funzione nel piano $x = x_0$:

$$\begin{cases} z = f(x_0, y_0) + \frac{\delta f}{\delta y}(x_0, y_0) \cdot (y - y_0) \\ x = x_0 \end{cases}$$

3. costruiamo il piano che contiene entrambe le rette:

$$z = f(x_0, y_0) + \frac{\delta f}{\delta x}(x_0, y_0) \cdot (x - x_0) + \frac{\delta f}{\delta y}(x_0, y_0) \cdot (y - y_0)$$

Il procedimento appena mostrato individua il piano tangente nell'ipotesi che esso esista, potrebbe però non esserci.

Differenziabilità e approssimazione lineare

In due o più variabili la sola derivabilità non implica né continuità né l'esistenza del piano tangente. Concetto di differenziabilità in più variabili: l'incremento di f è uguale all'incremento calcolato lungo il piano tangente, più un infinitesimo di ordine superiore rispetto alla lunghezza dell'incremento (h, k) delle variabili indipendenti. In formule:

$$f(x_0 + h, y_0 + k) - f(x_0, y_0) = \frac{\delta f}{\delta x}(x_0, y_0) \cdot (x - x_0) + \frac{\delta f}{\delta y}(x_0, y_0) \cdot (y - y_0) + o(\sqrt{h^2 + k^2})$$

per $(h, k) \rightarrow (0, 0)$.

Tutto ciò che è prima dell'uguale (primo membro) rappresenta l'incremento della funzione, i primi due addendi del secondo membro rappresentano l'incremento calcolato lungo il piano tangente. Ricordiamo che l'ultimo addendo rappresenta una funzione tale che $\lim_{(h,k) \rightarrow (0,0)} \frac{o(\sqrt{h^2+k^2})}{\sqrt{h^2+k^2}} = 0$. Se l'equazione di prima è soddisfatta, diremo che la funzione è differenziabile in (x_0, y_0) .

Da notare che la differenziabilità implica la derivabilità, cioè se una funzione è differenziabile in un punto, allora è anche derivabile nello stesso.

Se f è differenziabile in x_0 , si dice differenziale di f calcolato in x_0 la funzione lineare $df(x_0) : \mathbb{R}^n \rightarrow \mathbb{R}$ definita da:

$$df(x_0) : h \rightarrow \nabla f(x_0) \cdot h.$$

Nel caso in due variabili, il numero $\nabla f(x_0) \cdot h$ rappresenta l'incremento della funzione nel passare da x_0 a $x_0 + h$, calcolato lungo il piano tangente al grafico di f in x_0 .

L'approssimazione dell'incremento di f con il suo differenziale prende il nome di linearizzazione.

Verifica della differenziabilità

Per dimostrare la differenziabilità in un punto (x_0, y_0) bisogna provare che:

$$\lim_{h,k \rightarrow 0,0} \frac{f(x_0 + h, y_0 + k) - \{f(x_0, y_0) + \frac{\delta f}{\delta x}(x_0, y_0)h + \frac{\delta f}{\delta y}(x_0, y_0)k\}}{\sqrt{h^2 + k^2}} = 0.$$

dove $h = x - x_0$ e $k = y - y_0$.

Ma per certi casi particolari esistono criteri molto comodi e più semplici.

Teorema di condizione sufficiente di differenziabilità: se le derivate parziali di f esistono in un intorno di x_0 e sono continue in x_0 , allora f è differenziabile in x_0 .

In particolare se le derivate parziali esistono e sono continue in tutto A , allora f è differenziabile in tutto A .

Una funzione le cui derivate parziali esistono e sono continue in tutto A si dice di classe $C^1(A)$, dunque: $f \in C^1(A) \rightarrow f$ differenziabile in A .

Negli esercizi spesso si usa anche l'omogeneità di una funzione per sapere se essa è differenziabile o continua, oppure le proprietà delle funzioni radiali.

Negli esercizi seguire quest'ordine:

- E' continua nel punto richiesto? se non lo è, può essere allungata?
 - Funzione radiale? (vedi più avanti)
 - Funzione omogenea? (vedi più avanti)
 - Calcolo delle derivate parziali nel punto. Sono continue in quel punto?
 - Verifica della differenziabilità tramite la definizione.
-

Derivate direzionali

Si dice derivata direzionale della funzione f rispetto al versore v , nel punto x_0 , il limite

$$D_v f(x_0) = \lim_{t \rightarrow 0} \frac{f(x_0 + tv) - f(x_0)}{t}$$

purchè esista finito.

Detto in maniera diversa significa considerare la restrizione della funzione f alla direzione della retta passante per x_0 con direzione v , cioè $g(t) = f(x_0 + tv)$, e calcolarne la derivata, cioè $D_v f(x_0) = g'(0)$.

Calcolo di una derivata direzionale per un generico vettore $(\cos(\theta), \sin(\theta))$ nell'origine di una funzione f : Per prima cosa si ottiene la funzione $g(t) = f(t \cdot \cos(\theta), t \cdot \sin(\theta))$ e la si semplifica per $t \rightarrow 0$ (anche usando asintotici). In seguito si studia la derivata $g'(0) = \frac{\delta f}{\delta t}(t \cdot \cos(\theta), t \cdot \sin(\theta))$.

Se è richiesto il calcolo in un punto generico, e non nell'origine, è sufficiente usare $t \cdot \cos(\theta) + x_0$ e $t \cdot \sin(\theta) + y_0$.

Formula del gradiente: $D_v f(x_0) = \nabla f(x_0) \cdot v = \sum_{i=1}^n \frac{\delta f}{\delta x_i}(x_0) \cdot v_i$. Cioè la derivata direzionale è il prodotto scalare del gradiente con il versore nella direzione in cui si deriva, quindi tutte le derivate direzionali sono combinazioni lineari delle derivate parziali. Nel caso in due variabili la formula si riduce a $D_v f(x_0) = \nabla f(x_0, y_0) \cdot v = \frac{\delta f}{\delta x}(x_0, y_0)\cos(\theta) + \frac{\delta f}{\delta y}(x_0, y_0)\sin(\theta)$.

Se la formula del gradiente non vale in un punto, allora la funzione non è differenziabile in quel punto. Inoltre la formula del gradiente non vale se la generica derivata direzionale non è combinazione lineare di $\cos(\theta)$, $\sin(\theta)$.

Da notare è che $\nabla f(x_0)$ indica la direzione di massima crescita di f , ossia la direzione di massima derivata direzionale, invece $-\nabla f(x_0)$ rappresenta la direzione di minima derivata direzionale, infine nelle direzioni ortogonali al gradiente le derivate direzionali sono nulle.

Riepilogo

- $f \in C^1(A) \Rightarrow f$ differenziabile in A (cioè f ha iperpiano tangente) $\Rightarrow f$ è continua, derivabile, ha derivate direzionali, vale la formula del gradiente.
- f continua, derivabile, dotata di tutte le derivate direzionali $\nRightarrow f$ differenziabile
- f derivabile, dotata di tutte le derivate direzionali $\nRightarrow f$ continua

Calcolo delle derivate

$$\begin{aligned}\delta(\alpha \cdot f + \beta \cdot g) &= \alpha \delta(f) + \beta \delta(g) \\ \delta(f \cdot g) &= g \cdot \delta(f) + f \cdot \delta(g) \\ \delta\left(\frac{f}{g}\right) &= \frac{g \cdot \delta(f) - f \delta(g)}{g^2} \\ h(x) = f(g(x)) = g \circ f &\Rightarrow h'(x) = f'(g(x))g'(x) \\ \frac{\delta}{\delta x}[|x|] &= \frac{|x|}{x}\end{aligned}$$

Per calcolare il valore di una derivata parziale in punto (x_0, y_0) secondo la definizione seguire questo procedimento. Se si richiede di calcolare il valore della derivata parziale di x (cioè $\frac{\delta f}{\delta x}(x_0, y_0)$), si parte dalla funzione $f(x, y)$ e si sostituisce $y = y_0$, ottenendo quindi $f(x, y_0)$, successivamente si calcola la derivata parziale, ottenendo dunque $\frac{\delta f}{\delta x}(x, y_0)$. Come ultima cosa si sostituisce $x = x_0$ e si arriva a un risultato numerico. Per la trovare il valore della derivata parziale di y in un preciso punto seguire lo stesso procedimento opposto.

In alcuni esercizi è richiesto di calcolare le derivate parziali in tutti i punti in cui esistono. Il procedimento tipico consiste nel calcolare per prima cosa le derivate parziali generiche. Una volta calcolare sapremo che sicuramente esistono dove queste sono definite (dominio), ma non siamo sicuri dei punti in cui non lo sono (al di fuori del dominio). Quindi dobbiamo analizzare singolarmente tutti i punti al di fuori del dominio e per farlo sfruttiamo il procedimento visto sopra, calcolando esplicitamente le derivate nei punti richiesti. Finché si tratta per esempio di calcolarle per un punto preciso non ci sono problemi, il calcolo è facile, ma ci sono alcuni casi difficili, per esempio:

- Calcolare le derivate parziali secondo la definizione lungo una retta. Per esempio in $y = 0$, per calcolare la $\frac{\delta f}{\delta x}(x_0, 0)$ non ci sono problemi, si procede come al solito. Ma per la $\frac{\delta f}{\delta y}(x_0, 0)$ ci sono difficoltà, siccome non possiamo sostituire le y con 0 e poi derivare per la y , dobbiamo ragionare così: la derivata non esiste a meno che non ci sia un valore che le x possono assumere che annullino la funzione (per gli es che ho fatto fino ad ora sono solo al numeratore). Il concetto generale è che se non si trovano valori per x_0 tali che annullino la funzione e quindi ci permettano di calcolare la derivata parziale, si finisce per tornare a guardare la derivata parziale generica e quindi a non trovarla per quella retta. [spiegato davvero male, ma è un concetto strano].

Per stabilire dove la funzione sia derivabile bisogna calcolare le derivate parziali e osservarne il dominio.

(n.b. tipicamente negli esercizi le funzioni sono descritte da un sistema che contiene una funzione prolungata nell'origine, in questo caso bisogna calcolare le derivate parziali al di fuori dell'origine e studiarne il dominio, in seguito bisogna calcolare il valore della derivata parziale nel punto $(0, 0)$ col metodo descritto precedentemente).

Gradiente di una funzione radiale

Si chiama funzione radiale una funzione h che dipende solo dalla distanza di dall'origine, ossia

$$h(x) = g(|x|).$$

ponendo $\rho = |x| = \sqrt{\sum_{j=1}^n x_j^2}$ si ha:

$$\begin{aligned}\nabla_\rho &= \left(\frac{x_1}{\rho}, \frac{x_2}{\rho}, \dots, \frac{x_n}{\rho}\right). \\ \nabla h(x) &= g'(|x|) \left(\frac{x_1}{|x|}, \dots, \frac{x_n}{|x|}\right) \\ |\nabla h(x)| &= |g'(|x|)|\end{aligned}$$

Le funzioni radiali sono spesso utilizzate negli esercizi in cui le incognite compaiono solo all'interno del termine $\sqrt{\sum_{j=1}^n x_j^2}$, in tal caso si ottiene $g(\rho)$ sostituendo ogni $\sqrt{\sum_{j=1}^n x_j^2}$ con ρ , successivamente si può procedere sfruttando le proprietà di continuità e differenziabilità delle funzioni radiali.

Criterio di continuità e differenziabilità per funzioni radiali

Sia $f : \mathbb{R}^n - \{0\} \rightarrow \mathbb{R}$ una funzione radiale, cioè $f(x) = g(|x|)$ con $g : (0, +\infty) \rightarrow \mathbb{R}$ e sia f continua fuori dall'origine. Allora:

- f è continua in 0, se e solo se esiste finito $\lim_{\rho \rightarrow 0^+} g(\rho)$;
- f è differenziabile in 0 se e solo se esiste $g'(0) = 0$.

Negli esercizi spesso si controlla prima la continuità nell'origine, se non lo è si allunga la funzione e successivamente si calcola la differenziabilità nell'origine.

Funzioni omogenee

Una funzione $f : \mathbb{R}^n \rightarrow \mathbb{R}$ (eventualmente definita solo per $x \neq 0$), non identicamente nulla, si dice positivamente omogenea di grado $\alpha \in \mathbb{R}$ se

$$f(\lambda x) = \lambda^\alpha f(x) \quad \forall x \in \mathbb{R}^n, x \neq 0, \lambda > 0.$$

La funzione f si dice omogenea di grado α se la formula di prima vale anche per $\lambda < 0$.

Se f è positivamente omogenea vale

$$f(x) = f(|x| \cdot \frac{x}{|x|}) = |x|^\alpha f(\frac{x}{|x|}).$$

In particolare se f è omogenea (o positivamente omogenea) di grado zero, significa che è costante su ogni retta (o semiretta) uscente dall'origine. Infatti, indicata con

$$r(t) = tv$$

con v vettore fissato, sarà

$$f(r(t)) = f(tv) = t^0 f(v) = f(v) = \text{costante}.$$

Più in generale, per una funzione in due variabili positivamente omogenea di grado α vale la seguente rappresentazione in coordinate polari:

$$f(\rho, \theta) = \rho^\alpha g(1, \theta)$$

per qualche $\alpha \in \mathbb{R}$ e qualche funzione $g : [0, 2\pi) \rightarrow \mathbb{R}$.

Sia $f : \mathbb{R}^n \rightarrow \mathbb{R}$ una funzione positivamente omogenea di grado α , definita e continua per $x \neq 0$. Allora:

- f è continua anche nell'origine se $\alpha > 0$; in questo caso $f(0) = 0$; f è discontinua nell'origine se $\alpha < 0$; è discontinua anche se $\alpha = 0$, tranne il caso banale in cui f è costante.
 - f è differenziabile nell'origine se $\alpha > 1$; non è differenziabile nell'origine se $\alpha < 1$, tranne il caso banale in cui $\alpha = 0$ e f è costante; se $\alpha = 1$, f è differenziabile se e solo se è una funzione lineare, (ossia $f(x) = a \cdot x$ per qualche vettore costante $a \in \mathbb{R}^n$).
-

Ortogonalità del gradiente con le curve di livello

Il gradiente è ortogonale a ogni punto alle linee di livello.

Equazione del trasporto

Si definisce equazione del trasporto la seguente:

$$c \frac{\delta u}{\delta x} + \frac{\delta u}{\delta t} = 0 \quad (?)$$

Teorema del valor medio. Sia $A \subset \mathbb{R}^n$ un aperto e $f : A \rightarrow \mathbb{R}$ una funzione differenziabile in A . Allora per ogni coppia di punti $x_0, x_1 \in A$, esiste un punto x^* tale per cui:

$$f(x_1) - f(x_0) = \nabla f(x^*) \cdot (x_1 - x_0).$$

In particolare:

$$|f(x_1) - f(x_0)| \leq |\nabla f(x^*)| \cdot |x_1 - x_0|.$$

Derivate di ordine superiore e approssimazioni successive

Derivate di ordine superiore

Teorema di Schwartz. Sia $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ con A aperto. Supponiamo che (per certi indici $i, j \in \{1, 2, \dots, n\}$) le derivate seconde miste $f_{x_i x_j}$ e $f_{x_j x_i}$ esistano in un certo x_0 e siano continue in x_0 ; allora esse coincidono in x_0 .

Una funzione che ha tutte le derivate parziali seconde continue in un aperto A si dice di classe $C^2(A)$.

Se $f \in C^2(A)$, allora $f \in C^1(A)$ (in particolare f è differenziabile), le derivate parziali prima sono differenziabili, le derivate parziali seconde sono continue, le derivate seconde miste sono uguali.

Differenziale secondo, matrice hessiana, formula di Taylor al secondo ordine

Se $f \in C^2(A)$ e $x_0 \in A$, si dice differenziale secondo di f in x_0 la funzione

$$d^2 f(x_0) : h \rightarrow \sum_{i=1}^n \sum_{j=1}^n \frac{\delta^2 f}{\delta x_i \delta x_j}(x_0) h_i h_j.$$

I vari coefficienti $\frac{\delta^2 f}{\delta x_i \delta x_j}(x_0)$ possono essere ordinati in una matrice detta Hessiana:

$$H_f(x_0) = \begin{pmatrix} f_{x_1 x_1}(x_0) & f_{x_1 x_2}(x_0) & \cdots & f_{x_1 x_n}(x_0) \\ f_{x_2 x_1}(x_0) & f_{x_2 x_2}(x_0) & \cdots & f_{x_2 x_n}(x_0) \\ \cdots & \cdots & \cdots & \cdots \\ f_{x_n x_1}(x_0) & f_{x_n x_2}(x_0) & \cdots & f_{x_n x_n}(x_0) \end{pmatrix}$$

In particolare, per due variabili:

$$H_f(x_0, y_0) = \begin{pmatrix} f_{xx}(x_0, y_0) & f_{xy}(x_0, y_0) \\ f_{yx}(x_0, y_0) & f_{yy}(x_0, y_0) \end{pmatrix}$$

Se f è di classe C^2 , la matrice Hessiana è simmetrica.

Formula di Taylor (resto secondo Lagrange). Sia $f \in C^2(A)$; per ogni $x_0 \in A$ e $h \in \mathbb{R}^n$, tale che $x_0 + h \in A$, esiste un numero reale $\delta \in (0, 1)$, dipendente da x_0 e h , tale che:

$$f(x_0 + h) = f(x_0) + \sum_{i=1}^n \frac{\delta f}{\delta x_i}(x_0) h_i + \frac{1}{2} \sum_{i,j=1}^n \frac{\delta^2 f}{\delta x_i \delta x_j}(x_0 + \delta h) h_i h_j.$$

Formula di Taylor (resto secondo Peano). Sia $f \in C^2(a)$. Per ogni $x_0 \in A$ vale la formula:

$$f(x_0 + h) = f(x_0) + \sum_{i=1}^n \frac{\delta f}{\delta x_i}(x_0) h_i + \frac{1}{2} \sum_{i,j=1}^n \frac{\delta^2 f}{\delta x_i \delta x_j}(x_0) h_i h_j + o(|h|^2).$$

Ottimizzazione, estremi liberi

Generalità sui problemi di ottimizzazione

- x_0 è detto punto di massimo (minimo) globale se per ogni x si ha $f(x) \leq f(x_0)$ ($f(x) \geq f(x_0)$);
- x_0 è detto punto di massimo (minimo) locale se esiste un intorno di x_0 detto U tale per cui per ogni $x \in U$ si ha $f(x) \leq f(x_0)$ ($f(x) \geq f(x_0)$).

Estremi liberi, condizioni necessarie del prim'ordine

Teorema di Fermat. Sia $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$, con A aperto e $x_0 \in A$ un punto di massimo o minimo locale per f . Se f è derivabile in x_0 , allora $\nabla f(x_0) = 0$.

I punti in cui il gradiente di una funzione si annulla si dicono punti critici o stazionari di f . Una volta individuati tutti i punti stazionari, si può iniziare un'analisi su di essi per verificare se sono o meno punti di massimo o minimo. Se non lo sono essi prendono il nome di punti di sella o colle. Da notare particolarmente è che una funzione può assumere valori di massimo o minimo anche in punti in cui non è derivabile, dunque questi punti vanno analizzati separatamente.

Forme quadratiche, classificazione

Un modo per determinare la natura di un punto stazionario è quello di analizzare il segno dell'incremento $\nabla f(x_0) = f(x_0 + h) - f(x_0)$. Se infatti si riesce a stabilire che $\nabla f(x_0)$ si mantiene di segno positivo o negativo, per ogni h di modulo abbastanza piccolo, possiamo dedurre che x_0 è punto di minimo o massimo locale. Se invece al variare di h , $\nabla f(x_0)$ cambia segno, siamo in presenza di un punto di sella.

Lo studio del segno di $\nabla f(x_0)$ riconduce all'analisi del segno del polinomio omogeneo di secondo grado nelle componenti di h (che prende il nome di forma quadratica) dato da

$$\sum_{i,j=1}^n f_{x_i x_j}(x_0) h_i h_j.$$

Ogni forma quadratica risulta associata a una matrice simmetrica M . Nel caso del differenziale la matrice M coincide con la matrice Hessiana.

Il segno della forma quadratica è quindi studiabile analizzando la sua matrice $M = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ con $a \neq 0$ nel seguente modo:

- è definitivamente positiva (negativa) se e solo se $\det(M) > 0$ e $a > 0$ ($a < 0$);
- indefinita se $\det(M) < 0$;
- semidefinita positiva (negativa) se e solo se $\det(M) = 0$ e $a > 0$ ($a < 0$).

Se $a = 0$ e $c \neq 0$, nelle affermazioni precedenti occorre sostituire a con c .

Forme quadratiche, test degli autovalori

Un importante test per determinare il segno di una funzione quadratica in \mathbb{R}^n è basato sul segno degli autovalori della matrice M .

Ricordiamo che un numero complesso λ e un vettore non nullo $v \in \mathbb{C}^n$ si dicono, rispettivamente, autovalore e autovettore (di λ) di una matrice M di ordine n , se soddisfano la relazione:

$$Mv = \lambda v$$

oppure

$$(M - \lambda I_n)v = 0.$$

Quest'ultima equazione ha soluzioni v non nulle se e solo se la matrice dei coefficienti è singolare, ovvero se λ è soluzione dell'equazione caratteristica:

$$\det(M - \lambda I) = 0$$

esistono esattamente n autovalori di M ciascuno contato secondo la propria molteplicità.

Le matrici M simmetriche hanno proprietà importanti:

- gli autovalori di M sono reali e possiedono autovettori reali;
- esistono n autovettori lineari che costituiscono una base ortonormale in \mathbb{R}^n ;
- La matrice $S = w_1, w_2, \dots, w_n$ le cui colonne sono gli autovettori lineari è ortogonale e diagonalizza M , precisamente:

$$S^T M S = \Lambda = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

Tornando allo studio del segno della forma quadratica con la sua matrice M :

- definitivamente positiva (negativa) se e solo se tutti gli autovalori di M sono positivi (negativi);
- semidefinita positiva (negativa) se e solo se tutti gli autovalori di M sono ≥ 0 (≤ 0) e almeno uno di essi è nullo;
- indefinita se M ha almeno un autovalore positivo e uno negativo.

Da notare è che per una forma quadratica l'origine è sempre un punto stazionario.

Studio della natura dei punti critici

Per estrarre informazioni su un punto critico x_0 occorre studiare e classificare la forma quadratica

$$q(h) = \sum_{i,j=1}^n f_{x_i x_j}(x_0) h_i h_j = h^T H_f(x_0) h$$

dove $H_f(x_0)$ è la matrice Hessiana di f in x_0 .

- se la forma quadratica è definitivamente positiva (negativa), allora x_0 è un punto di minimo (massimo) locale forte;
- se la forma quadratica è indefinita, allora x_0 è un punto di sella;
- se la forma quadratica è in x_0 semidefinita positiva (negativa) e non nulla, allora x_0 è di minimo (massimo) debole oppure di colle; la situazione cambia se la forma quadratica è semidefinita positiva (negativa) non solo in x_0 , ma anche per ogni x in un intorno di x_0 , f è convessa (concava), dunque un punto di minimo (massimo) debole;
- se la forma quadratica è nulla, allora non possiamo estrarne informazioni significanti.

Vediamo una strategia da seguire:

1. si isolano i punti di f che non sono regolari (es. non derivabili una o due volte). Questi punti dovranno essere analizzati separatamente;
2. trovare i punti critici risolvendo:

$$\begin{cases} f_{x_1}(x_1, x_2, \dots, x_n) = 0 \\ f_{x_2}(x_1, x_2, \dots, x_n) = 0 \\ \dots \\ f_{x_n}(x_1, x_2, \dots, x_n) = 0 \end{cases}$$

3. si studia il segno della forma quadratica per ogni punto critico, se è definita o indefinita si giunge a una conclusione con le regole dette precedentemente, se è nulla o semidefinita si ricorre a uno studio diretto di $\nabla f(x_0)$ in un intorno di x_0 .

Più precisamente, nel caso bidimensionale, per ogni punto critico:

1. si calcola l'Hessiana:

$$H_f(x_0, y_0) = \begin{pmatrix} f_{xx}(x_0, y_0) & f_{xy}(x_0, y_0) \\ f_{yx}(x_0, y_0) & f_{yy}(x_0, y_0) \end{pmatrix}$$

2. se $\det(H_f(x_0, y_0)) > 0$ e

- $f_{xx}(x_0, y_0) > 0$ allora (x_0, y_0) è di minimo locale forte;
- $f_{xx}(x_0, y_0) < 0$ allora (x_0, y_0) è di massimo locale forte;

(si noti che in questo caso $f_{xx}(x_0, y_0)$ e $f_{yy}(x_0, y_0)$ hanno lo stesso segno).

3. se $\det(H_f(x_0, y_0)) < 0$ allora (x_0, y_0) è punto di sella;

4. se $\det(H_f(x_0, y_0)) = 0$ occorre un'analisi ulteriore.

Funzioni convesse di n variabili

Generalità sulle funzioni convesse

Un insieme $\Omega \subseteq \mathbb{R}^n$ si dice convesso se per ogni coppia di punti $x_1, x_2 \in \Omega$ si ha che $[x_1, x_2] \subseteq \Omega$ (dove col simbolo $[x_1, x_2]$ si denota il segmento con estremi x_1, x_2); si dice strettamente convesso se per ogni coppia di punti $x_1, x_2 \in \Omega$ il segmento (x_1, x_2) privato degli estremi è strettamente contenuto in Ω .

Si dice epigrafico di una funzione $f : \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ l'insieme

$$\text{epi}(f) = \{(x, z) \in \mathbb{R}^{n+1} : z \geq f(x), x \in \Omega\}$$

Si dice che una funzione è convessa se $\text{epi}(f)$ è un sottoinsieme convesso, si dice che una funzione è concava se $-f$ è convessa.

Formalmente si dice che una funzione è convessa se e solo se per ogni $x_1, x_2, t \in [0, 1]$ vale la condizione

$$f(tx_2 + (1-t)x_1) \leq tf(x_2) + (1-t)f(x_1).$$

Si noti che $tx_2 + (1-t)x_1$ percorre il segmento $[x_1, x_2]$ al variare di $t \in [0, 1]$

Se f è convessa allora:

- f è continua;
- f ha derivate parziali destre e sinistre in ogni punto;
- nei punti in cui è derivabile, f è anche diverenziabile.

Teorema di convessità e piano tangente. Sia $f : \Omega \rightarrow \mathbb{R}$ differenziabile in Ω . Allora f è convessa in Ω se e solo se per ogni coppia di punti $x_0, x \in \Omega$ si ha:

$$f(x) \geq f(x_0) + \nabla f(x_0) \cdot (x - x_0).$$

In due dimensioni:

$$f(x, y) \geq f(x_0, y_0) + \frac{\delta f}{\delta x}(x_0, y_0)(x - x_0) + \frac{\delta f}{\delta y}(x_0, y_0)(y - y_0).$$

che geometricamente significa che il piano tangente in x_0, y_0 sta sotto f .

Teorema di convessità e matrice Hessiana. Sia $f \in C^2(\Omega)$, con Ω aperto convesso in \mathbb{R}^n . Se per ogni x_0 in Ω la forma quadratica $d^2f(x_0)$ è semidefinita positiva, allora f è convessa in Ω .

Ottimizzazione di funzioni convesse e concave

Nelle funzioni convesse (concave) i punti stazionari, se esistono, rappresentano minimi (massimi) globali. Inoltre se la funzione è strettamente convessa (concava), il punto critico è di minimo (massimo) globale forte, quindi, in particolare, è unico.

Funzioni definite implicitamente

Funzione implicita di una variabile

Teorema di Dini della funzione implicita. Sia A un aperto in \mathbb{R}^2 e $f : A \rightarrow \mathbb{R}$ una funzione $C^1(A)$. Supponiamo che in un punto $(x_0, y_0) \in A$ sia:

$$f(x_0, y_0) = 0 \quad e \quad f_y(x_0, y_0) \neq 0.$$

Allora esiste un intorno I di x_0 in \mathbb{R} e un'unica funzione $g : I \rightarrow \mathbb{R}$, tale che $y_0 = g(x_0)$ e

$$f(x, g(x)) = 0 \quad \forall x \in I.$$

Inoltre, $g \in C^1(I)$ e

$$g'(x) = -\frac{f_x(x, g(x))}{f_y(x, g(x))} \quad \forall x \in I.$$

Notiamo che se $f(x_0, y_0) = 0$ e $f_y(x_0, y_0) = 0$, ma $f_x(x_0, y_0) \neq 0$, il teorema è ancora applicabile scambiando gli ruoli di x e y .

In sostanza i punti in cui il teorema di Dini non è applicabile sono quelli in cui il gradiente di f si annulla, ossia i punti critici.

Complementi

Topologia e funzioni continue

Teorema di Bolzano-Weierstrass in \mathbb{R}^n . Sia $\{x_k\}_{k=1}^{\infty}$ una successione limitata. Allora essa ammette una sottosuccessione convergente.

Successione di Cauchy in \mathbb{R}^n . Sia $\{x_k\}_{k=1}^{\infty}$ una successione in \mathbb{R}^n . Si dice che la successione soddisfa la condizione di Cauchy se

$$\forall \epsilon > 0 \quad \exists n_0 : \quad \forall h, k \geq n_0 \quad si \quad ha \quad |x_h - x_k| < \epsilon.$$

Teorema di completezza di \mathbb{R}^n . Se $\{x_k\}_{k=1}^{\infty}$ è una successione di Cauchy in \mathbb{R}^n , allora converge.

Teorema dell'uniforme continuità. Si dice che $f : \Omega \rightarrow \mathbb{R}$ è uniformemente continua in $\Omega \subseteq \mathbb{R}^n$ se:

$$\forall \epsilon > 0, \exists \delta > 0 : \forall x_1, x_2 \in \Omega, \text{ se } |x_1 - x_2| < \delta \text{ allora } |f(x_1) - f(x_2)| < \epsilon$$

Teorema di Cantor-Heine. Sia $K \subset \mathbb{R}^n$ un insieme chiuso e limitato e $f : K \rightarrow \mathbb{R}$ una funzione continua. Allora f è uniformemente continua in K .

Sia $\Omega \subset \mathbb{R}^n$ un aperto e sia $f : \Omega \rightarrow \mathbb{R}$ uniformemente continua in Ω . Allora f è prolungabile con continuità fino alla frontiera di Ω , ossia esiste una funzione $\bar{f} : \bar{\Omega} \rightarrow \mathbb{R}$ continua in $\bar{\Omega}$ e tale che in Ω coincide con f .
