Analisi Matematica B
– *Lezione 10* –

Lorenzo Brasco

(Università degli Studi di Ferrara)

Ferrara, 1 Aprile 2020

III.5 Derivate successive

Abbiamo visto che una funzione f di N variabili si dice **derivabile** se esistono le derivate parziali

$$\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_N}$$

Si dirà che f è **derivabile due volte** se ognuna delle sue derivate parziali è a sua volta derivabile

In tal caso, per ogni $i = 1, \dots, N$, risultano quindi definite le derivate

$$\frac{\partial}{\partial x_1}\frac{\partial f}{\partial x_i}, \; \frac{\partial}{\partial x_2}\frac{\partial f}{\partial x_i}, \; \dots, \; \frac{\partial}{\partial x_N}\frac{\partial f}{\partial x_i}$$

Notazione

Indichiamo con il simbolo

$$\boxed{\frac{\partial^2 f}{\partial x_j \, \partial x_i}} \qquad \text{la derivata parziale} \qquad \frac{\partial}{\partial x_j}\frac{\partial f}{\partial x_i}$$

Esempio

Calcoliamo le derivate parziali seconde della funzione

$$f(x,y)=\cos(x+2\,y)$$

- Le derivate parziali prime sono date da

$$\frac{\partial f}{\partial x}=-\sin(x+2\,y) \qquad \text{e} \qquad \frac{\partial f}{\partial y}=-2\,\sin(x+2\,y)$$

- queste funzioni sono ancora derivabili ed abbiamo

$$\frac{\partial^2 f}{\partial x^2}=\frac{\partial}{\partial x}\frac{\partial f}{\partial x}=-\cos(x+2\,y)$$

$$\frac{\partial^2 f}{\partial y\,\partial x}=\frac{\partial}{\partial y}\frac{\partial f}{\partial x}=-2\,\cos(x+2\,y)$$

$$\frac{\partial^2 f}{\partial x\,\partial y} = \frac{\partial}{\partial x}\frac{\partial f}{\partial y} = -2\,\cos(x + 2\,y)$$

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

- Si noti in particolare che una funzione di 2 variabili ha $2 \times 2 = 4$ derivate parziali seconde

Osservazione

Analogamente una funzione di 3 variabili ha $3 \times 3 = 9$ derivate parziali seconde

Più in generale, una funzione di N variabili ha $N \times N = N^2$ derivate parziali seconde

Vediamo un altro

Esempio

Calcoliamo le derivate parziali seconde della funzione

$$g(x, y) = x^2 - 3\,x\,y - y^2$$

- Le derivate parziali prime sono date da

$$\frac{\partial g}{\partial x} = 2\,x - 3\,y \qquad \text{e} \qquad \frac{\partial g}{\partial y} = -3\,x - 2\,y$$

- queste funzioni sono ancora derivabili ed abbiamo

$$\frac{\partial^2 g}{\partial x^2} = \frac{\partial}{\partial x}\frac{\partial g}{\partial x} = 2$$

$$\frac{\partial^2 g}{\partial y\,\partial x} = \frac{\partial}{\partial y}\frac{\partial g}{\partial x} = -3$$

$$\frac{\partial^2 g}{\partial x\,\partial y} = \frac{\partial}{\partial x}\frac{\partial g}{\partial y} = -3$$

$$\frac{\partial^2 g}{\partial y^2} = \frac{\partial}{\partial y}\frac{\partial g}{\partial y} = -2$$

Si osservi che anche stavolta abbiamo

$$\frac{\partial^2 g}{\partial y\,\partial x} = \frac{\partial^2 g}{\partial x\,\partial y}$$

sarà un caso....

Definizione

La matrice quadrata $N \times N$ composta da tutte le derivate parziali seconde

$$\begin{bmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1\,\partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_1\,\partial x_N} \\ \\ \dfrac{\partial^2 f}{\partial x_2\,\partial x_1} & \dfrac{\partial^2 f}{\partial x_2^2} & \cdots & \dfrac{\partial^2 f}{\partial x_2\,\partial x_N} \\ \\ \vdots & \vdots & \vdots & \vdots \\ \\ \dfrac{\partial^2 f}{\partial x_N\,\partial x_1} & \dfrac{\partial^2 f}{\partial x_N\,\partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_N^2} \end{bmatrix}$$

si chiama **matrice hessiana di** f. La indichiamo col simbolo Hf

Casi particolari

- Nel caso di $N = 2$ variabili, si ha

$$Hf(x,y) = \begin{bmatrix} \dfrac{\partial^2 f}{\partial x^2} & \dfrac{\partial^2 f}{\partial x\,\partial y} \\[2ex] \dfrac{\partial^2 f}{\partial y\,\partial x} & \dfrac{\partial^2 f}{\partial y^2} \end{bmatrix}$$

- per una funzione di $N = 3$ variabili si ha

$$Hf(x,y) = \begin{bmatrix} \dfrac{\partial^2 f}{\partial x^2} & \dfrac{\partial^2 f}{\partial x\,\partial y} & \dfrac{\partial^2 f}{\partial x\,\partial z} \\[2ex] \dfrac{\partial^2 f}{\partial y\,\partial x} & \dfrac{\partial^2 f}{\partial y^2} & \dfrac{\partial^2 f}{\partial y\,\partial z} \\[2ex] \dfrac{\partial^2 f}{\partial z\,\partial x} & \dfrac{\partial^2 f}{\partial z\,\partial y} & \dfrac{\partial^2 f}{\partial z^2} \end{bmatrix}$$

Vale il seguente importante risultato, di cui omettiamo la dimostrazione

Teorema (Schwarz)

Sia $A \subset \mathbb{R}^N$ aperto e sia $f : A \to \mathbb{R}$ una funzione C^2 su A, ovvero

f è derivabile due volte e tutte le sue derivate parziali di ordine 1 e 2 sono continue su A

Allora si ha

$$\frac{\partial^2 f}{\partial x_i \, \partial x_j} = \frac{\partial^2 f}{\partial x_j \, \partial x_i} \qquad \text{per ogni } i \neq j$$

In particolare, se f è C^2, la matrice hessiana Hf è **simmetrica**

III.6 Formula di Taylor di ordine 2

Iniziamo questa sezione ricordando che...

MEMO 5

Se $f : (a, b) \to \mathbb{R}$ è derivabile <u>due volte</u> in $x_0 \in (a, b)$ possiamo migliorare l'approssimazione asintotica con la retta tangente e scrivere

$$\begin{aligned} f(x) &= f(x_0) + f'(x_0)\,(x - x_0) \\ &\quad + \frac{1}{2}\,f''(x_0)\,(x - x_0)^2 + o\big((x - x_0)^2\big) \qquad \textit{se } x \to x_0 \end{aligned}$$

Formula di Taylor all'ordine 2 *centrata in* x_0, *con resto di Peano*

Osservazione

Vale qualcosa di analogo per le funzioni di più variabili...

Teorema (Formula di Taylor al secondo ordine)

Sia $A \subset \mathbb{R}^N$ *un aperto*

Sia $f : A \to \mathbb{R}$ *una funzione di classe* C^2 *su* A

Per ogni $\mathbf{x}_0 \in A$ *vale l'identità asintotica*

$$\begin{aligned} f(\mathbf{x}) = f(\mathbf{x}_0) &+ \langle \nabla f(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle \\ &+ \frac{1}{2} \langle Hf(\mathbf{x}_0)\,(\mathbf{x} - \mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle \\ &+ o\big(|\mathbf{x} - \mathbf{x}_0|^2\big), \qquad \text{se } \mathbf{x} \to \mathbf{x}_0 \end{aligned}$$

Attenzione

$Hf(\mathbf{x}_0)$ è una matrice $N \times N$ $\qquad$ $(\mathbf{x} - \mathbf{x}_0)$ è un vettore di $\mathbb{R}^N$

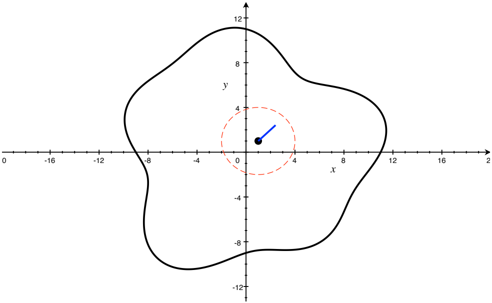
La scrittura $Hf(\mathbf{x}_0)\,(\mathbf{x} - \mathbf{x}_0)$ indica l'operazione di "prodotto matrice $N \times N$ per vettore **colonna**", che da' come risultato un vettore di $\mathbb{R}^N$

Dimostrazione

- La dimostrazione si basa fortemente sulla formula di Taylor per funzioni di una variabile
- dal momento che A è aperto, dato $\mathbf{x}_0 \in A$ esiste $B_R(\mathbf{x}_0) \subset A$
- fissiamo adesso una direzione $\mathbf{v}$ (ovvero $|\mathbf{v}| = 1$), allora avremo che

$$\mathbf{x}_0 + t\,\mathbf{v} \in B_R(\mathbf{x}_0) \subset A,$$

per t piccoli abbastanza



- introduciamo la funzione di una variabile

$$\psi(t) = f(\mathbf{x}_0 + t\,\mathbf{v})$$

- ovvero ψ è la composizione di f con la curva $\gamma(t) = \mathbf{x}_0 + t\,\mathbf{v}$ il cui sostegno è un segmento passante da $\mathbf{x}_0$, con direzione $\mathbf{v}$
- osserviamo che in base al *Teorema "Derivata curvilinea"*, si ha

$$\begin{aligned}\psi'(t) &= \langle \nabla f(\mathbf{x}_0 + t\,\mathbf{v}), \frac{d}{dt}(\mathbf{x}_0 + t\,\mathbf{v})\rangle \\ &= \langle \nabla f(\mathbf{x}_0 + t\,\mathbf{v}), \mathbf{v}\rangle \\ &= \sum_{i=1}^{N} \frac{\partial f}{\partial x_i}(\mathbf{x}_0 + t\,\mathbf{v})\, v_i\end{aligned}$$

da cui

$$\psi'(0) = \langle \nabla f(\mathbf{x}_0), \mathbf{v}\rangle$$

- sempre usando il *Teorema "Derivata curvilinea"* possiamo calcolare anche la derivata seconda

$$\begin{aligned}
\psi''(t) &= \frac{d}{dt}\left(\sum_{i=1}^{N} \frac{\partial f}{\partial x_i}(\mathbf{x}_0 + t\,\mathbf{v})\, v_i\right) \\
&= \sum_{i=1}^{N} \frac{d}{dt}\left(\frac{\partial f}{\partial x_i}(\mathbf{x}_0 + t\,\mathbf{v})\right)\, v_i \\
&= \sum_{i=1}^{N} \langle \nabla \frac{\partial f}{\partial x_i}(\mathbf{x}_0 + t\,\mathbf{v}), \mathbf{v}\rangle\, v_i
\end{aligned}$$

da cui

$$\begin{aligned}
\psi''(0) &= \sum_{i=1}^{N} \langle \nabla \frac{\partial f}{\partial x_i}(\mathbf{x}_0), \mathbf{v}\rangle\, v_i \\
&= \sum_{i=1}^{N}\left(\sum_{j=1}^{N} \frac{\partial^2 f}{\partial x_j\, \partial x_i}\, v_j\right)\, v_i = \langle Hf(\mathbf{x}_0)\,\mathbf{v}, \mathbf{v}\rangle
\end{aligned}$$

- ricapitolando, per la funzione di una variabile

$$\psi(t) = f(\mathbf{x}_0 + t\,\mathbf{v})$$

abbiamo

$$\boxed{\psi(0) = f(\mathbf{x}_0)} \qquad \boxed{\psi'(0) = \langle \nabla f(\mathbf{x}_0), \mathbf{v} \rangle}$$

e

$$\boxed{\psi''(0) = \langle Hf(\mathbf{x}_0)\,\mathbf{v}, \mathbf{v} \rangle}$$

- d'altra parte, dalla formula di Taylor al secondo ordine per funzioni di una variabile, vale

$$\psi(t) = \psi(0) + \psi'(0)\,t + \frac{1}{2}\,\psi''(0)\,t^2 + o(t^2)$$

- usando le identità nei riquadri, si conclude

Commento

La formula di Taylor di ordine 2 sarà di fondamentale importanza per capire se un punto critico $\mathbf{x}_0$, ovvero

$$\nabla f(\mathbf{x}_0) = \mathbf{0}$$

è un punto di massimo o minimo locale per f

- Ricordate che siamo interessati al

"Problema modello"

Date due funzioni continue f e g, determinare

$$\min \Big\{ f(\mathbf{x}) \, : \, g(\mathbf{x}) \leq 0 \Big\} \qquad \text{e} \qquad \max \Big\{ f(\mathbf{x}) \, : \, g(\mathbf{x}) \leq 0 \Big\}$$

Esercizio

Si dia lo sviluppo di Taylor al secondo ordine centrato in $(1,0)$ *per la funzione*

$$f(x,y) = x^3 + x\,y + y^2$$

Soluzione

- Si tratta di una funzione C^2 su tutto $\mathbb{R}^2$
- in base al Teorema precedente, ci servono

$$f(1,0), \qquad \nabla f(1,0), \qquad Hf(1,0)$$

- si ha $f(1,0) = 1$

- per il gradiente si ha

$$\nabla f(x, y) = (3\,x^2 + y, x + 2\,y)$$

da cui

$$\nabla f(1, 0) = (3, 1)$$

- la matrice hessiana è data da

$$Hf(x, y) = \begin{bmatrix} 6\,x & 1 \\ 1 & 2 \end{bmatrix}$$

da cui

$$Hf(1, 0) = \begin{bmatrix} 6 & 1 \\ 1 & 2 \end{bmatrix}$$

- in base alla formula generale

$$
\begin{aligned}
f(x,y) = f(1,0) &+ \langle \nabla f(1,0), (x,y)-(1,0)\rangle \\
&+ \frac{1}{2}\langle Hf(1,0)\big((x,y)-(1,0)\big), (x,y)-(1,0)\rangle \\
&+ o\Big(|(x,y)-(1,0)|^2\Big) \qquad \text{se } (x,y)\to(1,0)
\end{aligned}
$$

- ...troviamo

$$
\begin{aligned}
x^3 + x\,y + y^2 = 1 &+ \left\langle \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \begin{bmatrix} x-1 \\ y \end{bmatrix} \right\rangle \\
&+ \frac{1}{2}\left\langle \begin{bmatrix} 6 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x-1 \\ y \end{bmatrix}, \begin{bmatrix} x-1 \\ y \end{bmatrix} \right\rangle \\
&+ o\left((x-1)^2 + y^2\right)
\end{aligned}
$$

- sviluppando i conti a secondo membro

$$
\begin{aligned}
x^3 + x\,y + y^2 &= 1 + \left\langle \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \begin{bmatrix} x-1 \\ y \end{bmatrix} \right\rangle \\
&\quad + \frac{1}{2} \left\langle \begin{bmatrix} 6 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x-1 \\ y \end{bmatrix}, \begin{bmatrix} x-1 \\ y \end{bmatrix} \right\rangle \\
&\quad + o\left((x-1)^2 + y^2\right) \\
&= 1 + 3\,(x-1) + y \\
&\quad + \frac{1}{2} \left\langle \begin{bmatrix} 6\,(x-1) + y \\ (x-1) + 2\,y \end{bmatrix}, \begin{bmatrix} x-1 \\ y \end{bmatrix} \right\rangle \\
&\quad + o\left((x-1)^2 + y^2\right)
\end{aligned}
$$

- calcolando anche l'ultimo prodotto scalare, si ottiene infine...

- ...lo sviluppo di Taylor

$$
\begin{aligned}
x^3 + x\,y + y^2 = 1 &+ \boxed{3\,(x-1) + y} \\
&+ \boxed{3\,(x-1)^2 + y\,(x-1) + y^2} \\
&+ o\left(x^2 + (y-1)^2\right)
\end{aligned}
$$

Esercizio (per casa)

Si dia lo sviluppo di Taylor al secondo ordine centrato in $(0,1)$ *per la funzione*

$$f(x,y) = x^3 + x\,y + y^2$$

Capitolo IV

“*Classificazione dei punti critici*”

Preambolo

Ricordiamo che un **punto critico** $\mathbf{x}_0$ per una funzione f di N variabili è un punto per cui

$$\nabla f(\mathbf{x}_0) = \mathbf{0}$$

In base al *Teorema di Fermat*, i punti critici $\mathbf{x}_0 \in A$ con A aperto, sono candidati punti di minimo locale o massimo locale

Obiettivo

In questo capitolo vogliamo dare delle condizioni sufficienti per capire quando un punto critico è di minimo/massimo locale o nessuno dei due

IV.1 Un passo indietro: funzioni di una variabile

- Torniamo un attimo alle funzioni di una variabile
$$f : (a, b) \to \mathbb{R}$$
- Se f è derivabile su (a, b), i punti critici sono quei punti $x_0 \in (a, b)$ tali che
$$f'(x_0) = 0$$
- Anche qua, in base dal *Teorema di Fermat*, avevamo visto che questi punti sono candidati punti di minimo locale o massimo locale
- come facevamo per decidere la loro natura?

Lo strumento principale che avevamo usato era il

MEMO "Test di monotonia"

$$f'(x) \geq 0 \qquad \Longrightarrow \qquad f \text{ crescente}$$

$$f'(x) \leq 0 \qquad \Longrightarrow \qquad f \text{ decrescente}$$

Quindi se avevamo $f'(x_0) = 0$ e per esempio succedeva che

$$\begin{cases} f'(x) \geq 0 & \text{per } x \geq x_0 \\ f'(x) \leq 0 & \text{per } x \leq x_0 \end{cases}$$

potevamo dire che f era **punto di minimo locale**

Domanda

Possiamo generalizzare questo metodo a più variabili?

Risposta

Sfortunatamente NO

Come abbiamo visto, in più variabili ci sono infinite direzioni in cui guardare gli incrementi della funzione

In più variabili, l'applicazione del *Test di monotonia* implicherebbe di studiare il segno di tutte le derivate direzionali

$$\frac{\partial f}{\partial \mathbf{v}}$$

per punti vicini al punto critico $\mathbf{x}_0$

— Troppo complicato! —

D'altra parte, per funzioni di una variabile, c'era un altro metodo per studiare la natura di un punto critico...

...un metodo che fin'ora vi avevo tenuto segreto

Un metodo di classificazione "puntuale"

Sia $f : (a, b) \to \mathbb{R}$ *una funzione* **derivabile due volte**

Supponiamo che $x_0 \in (a, b)$ *sia tale che*

$$f'(x_0) = 0 \qquad e \qquad \boxed{f''(x_0) > 0}$$

Allora x_0 *è* **punto di minimo locale***!*

Se invece

$$f'(x_0) = 0 \qquad e \qquad \boxed{f''(x_0) < 0}$$

Allora x_0 *è* **punto di massimo locale***!*

Dimostrazione

- Facciamo il primo caso $f''(x_0) > 0$ (l'altro *per esercizio*)
- la dimostrazione si basa sulla *formula di Taylor all'ordine* 2, centrata nel punto critico x_0
- ovvero, sappiamo che vale

$$f(x) = f(x_0) + \cancel{f'(x_0)\,(x - x_0)} + \frac{f''(x_0)}{2}\,(x - x_0)^2 + o\big((x - x_0)^2\big), \qquad \text{se } x \to x_0$$

- il termine di ordine 1 l'abbiamo cancellato perché x_0 è **punto critico**
- abbiamo quindi

$$f(x) = f(x_0) + \frac{f''(x_0)}{2}\,(x - x_0)^2 + o\big((x - x_0)^2\big), \qquad \text{se } x \to x_0$$

- per assurdo, se x_0 non fosse di minimo locale, vorrebbe dire che esiste una successione $y_n \to x_0$ tale che

$$f(y_n) < f(x_0) \qquad \text{per ogni } n \in \mathbb{N}$$

- usando l'identità asintotica di ordine 2 precedente per i punti y_n, si ha quindi

$$f(x_0) > f(y_n) = f(x_0) + \frac{f''(x_0)}{2}(y_n - x_0)^2 + o\big((y_n - x_0)^2\big)$$

- ovvero

$$f(x_0) > f(x_0) + \frac{f''(x_0)}{2}(y_n - x_0)^2 + o\big((y_n - x_0)^2\big)$$

- eliminando il termine comune $f(x_0)$, si ha dunque

$$0 > \frac{f''(x_0)}{2}\,(y_n - x_0)^2 + o\big((y_n - x_0)^2\big)$$

- se si divide adesso per $(y_n - x_0)^2$ e si passa al limite, dalla definizione di $o-$piccolo si ha

$$0 \geq \frac{f''(x_0)}{2}$$

- questo contraddice l'ipotesi che $f''(x_0) > 0$....quindi x_0 è punto di minimo locale!

Buone notizie!

Anche per una funzione di più variabili, abbiamo la *formula di Taylor all'ordine* 2 (vedi *Capitolo III*)

$$\begin{aligned} f(\mathbf{x}) = f(\mathbf{x}_0) &+ \langle \nabla f(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle \\ &+ \frac{1}{2} \langle Hf(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle \\ &+ o(|\mathbf{x} - \mathbf{x}_0|^2) \qquad \text{se } \mathbf{x} \to \mathbf{x}_0 \end{aligned}$$

Possiamo quindi tentare di replicare la strategia precedente per funzioni di più variabili

Questa strada funzionerà, ogni volta che riusciamo a dire che la matrice simmetrica $Hf(\mathbf{x}_0)$ è tale che

$$\langle Hf(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle > 0$$

per ogni $\mathbf{x} \to \mathbf{x}_0$

Siamo sulla buona strada!

Quindi, tutto ciò che ci serve per far funzionare questo metodo in più variabili è capire

come si fa a dire che una matrice simmetrica è "positiva" o "negativa"?

IV.2 Matrici simmetriche

Questa sezione contiene fatti che dovrebbero essere noti da GEOMETRIA E ALGEBRA

Definizione

Sia $A \in \mathcal{M}_N(\mathbb{R})$ una matrice $N \times N$ simmetrica

Si dice che:

- A è **definita positiva** se
$$\langle A\mathbf{v}, \mathbf{v} \rangle > 0, \qquad \text{per ogni } \mathbf{v} \neq \mathbf{0}$$
- A è **definita negativa** se
$$\langle A\mathbf{v}, \mathbf{v} \rangle < 0, \qquad \text{per ogni } \mathbf{v} \neq \mathbf{0}$$
- A è **indefinita** se esistono $\mathbf{v}$ e $\mathbf{w}$ non nulli, tali che
$$\langle A\mathbf{v}, \mathbf{v} \rangle > 0 \qquad \text{e} \qquad \langle A\mathbf{w}, \mathbf{w} \rangle < 0$$

Per le matrici 2×2 simmetriche, abbiamo il seguente semplice criterio per riconoscerne la *segnatura*

Teorema (segnatura di una 2×2)

Sia $A \in \mathcal{M}_2(\mathbb{R})$ una matrice 2×2 simmetrica. Allora:

1. A è definita positiva se e solo se
$$\boxed{\det(A) > 0 \qquad \text{e} \qquad \mathrm{tr}(A) > 0}$$
2. A è definita negativa se e solo se
$$\boxed{\det(A) > 0 \qquad \text{e} \qquad \mathrm{tr}(A) < 0}$$
3. A è indefinita se e solo se
$$\boxed{\det(A) < 0}$$

Ricorda

$\mathrm{tr}(A)$ indica la **traccia**, ovvero la somma degli elementi diagonali

L'interesse delle matrici definite positive o negative, per quanto riguarda il problema della classificazione dei punti critici in più variabili, è contenuto nella seguente

Proposizione

Sia $A \in \mathcal{M}_N(\mathbb{R})$ una matrice $N \times N$ simmetrica. Allora:

- *se A è definita positiva, esiste una costante $c > 0$ tale che*

$$\langle A\,\mathbf{v}, \mathbf{v} \rangle \geq c\,|\mathbf{v}|^2, \qquad \text{per ogni } \mathbf{v} \in \mathbb{R}^N$$

- *se A è definita negativa, esiste una costante $c > 0$ tale che*

$$\langle A\,\mathbf{v}, \mathbf{v} \rangle \leq -c\,|\mathbf{v}|^2, \qquad \text{per ogni } \mathbf{v} \in \mathbb{R}^N$$

Dimostrazione

- Basta dimostrare il primo fatto
- il secondo ne segue facilmente, osservando che se A è definita negativa, allora $-A$ è definita positiva
- sia A definita positiva, ovvero

$$\langle A\,\mathbf{v}, \mathbf{v}\rangle > 0, \qquad \text{per ogni } \mathbf{v} \neq \mathbf{0}$$

- in particolare

$$\langle A\,\mathbf{v}, \mathbf{v}\rangle > 0, \qquad \text{per ogni } \mathbf{v} \text{ tale che } |\mathbf{v}| = 1 \tag{1}$$

- definiamo la **sfera**

$$S = \{\mathbf{v} \in \mathbb{R}^N \,:\, |\mathbf{v}| = 1\}$$

- consideriamo la funzione $h : S \to \mathbb{R}$ definita da

$$h(\mathbf{v}) = \langle A\,\mathbf{v}, \mathbf{v}\rangle, \qquad \text{per ogni } \mathbf{v} \in S$$

- S è un insieme chiuso e limitato, per definizione
- h è continua su S
- dal *Teorema di Weierstrass* esistono

$$\max_S h \qquad \text{e} \qquad \min_S h$$

- in particolare esiste $\mathbf{v}_0 \in S$ punto di minimo per la funzione h su S
- ovvero

$$h(\mathbf{v}) \geq h(\mathbf{v}_0), \qquad \text{per ogni } \mathbf{v} \in S.$$

- ma $\mathbf{v}_0 \in S$, ricordando (1) si ha

$$h(\mathbf{v}) \geq h(\mathbf{v}_0) > 0, \qquad \text{per ogni } \mathbf{v} \in S.$$

- definiamo $c = h(\mathbf{v}_0) > 0$
- usiamo la bilinearità del prodotto scalare (vedi *Lezione 1*), si ha per ogni $\mathbf{v} \neq \mathbf{0}$

$$\begin{aligned}\langle A\,\mathbf{v}, \mathbf{v}\rangle &= \left\langle A\,|\mathbf{v}|\,\frac{\mathbf{v}}{|\mathbf{v}|}, |\mathbf{v}|\,\frac{\mathbf{v}}{|\mathbf{v}|}\right\rangle \\ &= |\mathbf{v}|^2\,\left\langle A\,\frac{\mathbf{v}}{|\mathbf{v}|}, \frac{\mathbf{v}}{|\mathbf{v}|}\right\rangle \\ &= |\mathbf{v}|^2\,h\left(\frac{\mathbf{v}}{|\mathbf{v}|}\right) \geq |\mathbf{v}|^2\,h(\mathbf{v}_0) = c\,|\mathbf{v}^2|,\end{aligned}$$

  che conclude la dimostrazione
- si osservi che abbiamo usato che

$$\frac{\mathbf{v}}{|\mathbf{v}|} \in S \qquad \text{per ogni } \mathbf{v} \neq \mathbf{0}$$

IV.3 Criteri di classificazione

Prima di dare il criterio di classificazione dei punti critici, abbiamo bisogno della seguente

Definizione

Sia $A \subset \mathbb{R}^N$ aperto e sia $f : A \to \mathbb{R}$ derivabile

Un punto critico $\mathbf{x}_0 \in A$ si dice **punto sella per** f se non è ne' di massimo locale ne' di minimo locale

Osservazione

In altre parole, un punto sella ha le seguenti proprietà

- $\nabla f(\mathbf{x}_0) = \mathbf{0}$
- per ogni $R > 0$, esistono $\mathbf{x}_R, \mathbf{y}_R \in B_R(\mathbf{x}_0) \cap A$ tali che

$$f(\mathbf{x}_R) > f(\mathbf{x}_0) > f(\mathbf{y}_R)$$

Definizione

Sia $A \subset \mathbb{R}^N$ aperto e sia $f : A \to \mathbb{R}$ una funzione C^2 su A

Un punto critico $\mathbf{x}_0 \in A$ per f si dice

- **non degenere** se

$$\det Hf(\mathbf{x}_0) \neq 0$$

- **degenere** se

$$\det Hf(\mathbf{x}_0) = 0$$

Esempio

La funzione $f(x, y) = x^2 + y^2$ ha come unico punto critico $(0, 0)$

Esso è non degenere, dal momento che la sua matrice hessiana è costante

$$Hf(x, y) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

con determinante non nullo

Teorema

Sia $A \subset \mathbb{R}^N$ *aperto e sia* $f : A \to \mathbb{R}$ *una funzione* C^2 *su* A

Sia $\mathbf{x}_0 \in A$ *un punto critico non degenere di* f

1. *se*

   $$\boxed{Hf(\mathbf{x}_0) \text{ è definita positiva}}$$

   allora $\mathbf{x}_0$ è **punto di minimo locale**

2. *se*

   $$\boxed{Hf(\mathbf{x}_0) \text{ è definita negativa}}$$

   allora $\mathbf{x}_0$ è **punto di massimo locale**

3. *se*

   $$\boxed{Hf(\mathbf{x}_0) \text{ è indefinita}}$$

   allora $\mathbf{x}_0$ è **punto sella**

Dimostrazione

1. Supponiamo che $Hf(\mathbf{x}_0)$ sia definita positiva

- usiamo lo *sviluppo di Taylor all'ordine* 2 per f centrato nel punto critico $\mathbf{x}_0$

$$\begin{aligned} f(\mathbf{x}) = f(\mathbf{x}_0) &+ \langle \nabla f(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle \\ &+ \frac{1}{2} \langle Hf(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle \\ &+ o(|\mathbf{x} - \mathbf{x}_0|^2), \qquad \text{se } \mathbf{x} \to \mathbf{x}_0. \end{aligned}$$

- ma $\mathbf{x}_0$ è un punto critico, quindi $\nabla f(\mathbf{x}_0) = \mathbf{0}$ e dunque

$$\begin{gathered} f(\mathbf{x}) = f(\mathbf{x}_0) + \frac{1}{2} \langle Hf(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle \\ + o(|\mathbf{x} - \mathbf{x}_0|^2), \qquad \text{se } \mathbf{x} \to \mathbf{x}_0. \end{gathered}$$

- visto che $Hf(\mathbf{x}_0)$ è definita positiva, sappiamo che esiste $c > 0$ tale che

$$\langle Hf(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle \geq c\, |\mathbf{x} - \mathbf{x}_0|^2$$

- al momento, abbiamo quindi ottenuto che vale

$$f(\mathbf{x}) \geq f(\mathbf{x}_0) + \frac{c}{2}|\mathbf{x} - \mathbf{x}_0|^2 + o(|\mathbf{x} - \mathbf{x}_0|^2), \qquad \text{se } \mathbf{x} \to \mathbf{x}_0$$

- procediamo adesso <u>per assurdo</u>: se $\mathbf{x}_0$ non fosse punto di minimo locale, avremmo che in ogni palla $B_{\frac{1}{n}}(\mathbf{x}_0)$ esiste un punto $\mathbf{y}_n \neq \mathbf{x}_0$ tale che

$$f(\mathbf{y}_n) < f(\mathbf{x}_0)$$

- per costruzione, si ha $\mathbf{y}_n \to \mathbf{x}_0$, quindi dalla maggiorazione precedente

$$f(\mathbf{x}_0) > f(\mathbf{y}_n) \geq f(\mathbf{x}_0) + \frac{c}{2}|\mathbf{y}_n - \mathbf{x}_0|^2 + o(|\mathbf{y}_n - \mathbf{x}_0|^2), \qquad \text{se } n \to \infty$$

- otteniamo in particolare

$$\cancel{f(\mathbf{x}_0)} > \cancel{f(\mathbf{x}_0)} + \frac{c}{2}\,|\mathbf{y}_n - \mathbf{x}_0|^2 + o(|\mathbf{y}_n - \mathbf{x}_0|^2), \qquad \text{se } n \to \infty$$

- ...ovvero

$$\frac{c}{2}\,|\mathbf{y}_n - \mathbf{x}_0|^2 + o(|\mathbf{y}_n - \mathbf{x}_0|^2) < 0, \qquad \text{se } n \to \infty$$

- dividendo per $|\mathbf{y}_n - \mathbf{x}_0|^2$

$$\frac{c}{2} + \frac{o(|\mathbf{y}_n - \mathbf{x}_0|^2)}{|\mathbf{y}_n - \mathbf{x}_0|^2} < 0, \qquad \text{se } n \to \infty$$

- da cui passando al limite per $n \to \infty$ (ed usando la definizione di $o-$piccolo)

$$\frac{c}{2} \leq 0$$

- ricordando che $c > 0$ otteniamo un **assurdo**!

2. Supponiamo adesso che $Hf(\mathbf{x}_0)$ sia definita negativa
   - allora $-Hf(\mathbf{x}_0)$ è definita positiva
   - ma $-Hf(\mathbf{x}_0)$ è la matrice hessiana di $-f$
   - per il punto 1. $\mathbf{x}_0$ è punto di minimo locale per $-f$
   - ovvero $\mathbf{x}_0$ è punto di massimo locale per f