Massimi e minimi relativi per funzioni di più variabili

Sia $A \subseteq \mathbb{R}^N$ aperto, sia $\underline{x}_0 \in A$, sia $f: A \to \mathbb{R}$.

Definizione Diciamo che $\underline{x}_0$ è <u>punto di massimo relativo</u>, o <u>locale</u> per f se esiste $B(\underline{x}_0, R) \subseteq A$ tale che

$$f(\underline{x}) \leq f(\underline{x}_0) \qquad \forall\, \underline{x} \in B(\underline{x}_0, R).$$

Diciamo che $\underline{x}_0$ è <u>punto di minimo relativo</u>, o <u>locale</u>, per f se esiste $B(\underline{x}_0, R) \subseteq A$ tale che

$$f(\underline{x}) \geq f(\underline{x}_0) \qquad \forall\, \underline{x} \in B(\underline{x}_0, R).$$

Cosa succede nel caso di funzioni di una variabile? Se $f \in C^2[a,b]$ e $x_0 \in]a,b[$, si ha:

Proposizione Se x_0 è punto di massimo relativo per f, allora

$$f'(x_0) = 0, \quad f''(x_0) \leq 0;$$

se x_0 è punto di minimo relativo per f, allora

$$f'(x_0) = 0, \quad f''(x_0) \geq 0.$$

Viceversa, se

$$f'(x_0) = 0, \quad f''(x_0) < 0,$$

allora x_0 è punto di massimo relativo per f; se invece

$$f'(x_0) = 0, \quad f''(x_0) > 0,$$

allora x_0 è punto di minimo relativo per f.

Nel caso di N variabili si ha una generalizzazione naturale.

Anzitutto, se $f \in C^1(A)$, il punto $\underline{x}_0 \in A$ si dice punto stazionario per f se $\nabla f(\underline{x}_0) = \underline{0}$. La denominazione nasce dal fatto che se $\underline{x}_0$ è stazionario per f, allora il piano tangente al grafico di f in $(\underline{x}_0, f(\underline{x}_0))$ è orizzontale, con equazione $z = f(\underline{x}_0)$.

La generalizzazione è la seguente:

Teorema Sia $f \in C^2(A)$, sia $\underline{x}_0 \in A$.

(i) Se $\underline{x}_0$ è punto di massimo relativo per f, allora $\nabla f(\underline{x}_0) = \underline{0}$ e $H_f(\underline{x}_0)$ ha autovalori non positivi (≤ 0).

(ii) Se $\underline{x}_0$ è punto di minimo relativo per f, allora $\nabla f(\underline{x}_0) = \underline{0}$ e $H_f(\underline{x}_0)$ ha autovalori non negativi (≥ 0).

(iii) Se $\nabla f(\underline{x}_0) = \underline{0}$ e $H_f(\underline{x}_0)$ ha autovalori negativi, allora $\underline{x}_0$ è punto di massimo relativo per f.

(iv) Se $\nabla f(\underline{x}_0) = \underline{0}$ e $H_f(\underline{x}_0)$ ha autovalori positivi, allora $\underline{x}_0$ è punto di minimo relativo per f.

(v) Se $\nabla f(\underline{x}_0) = \underline{0}$ e $H_f(\underline{x}_0)$ ha sia autovalori positivi, sia autovalori negativi, allora $\underline{x}_0$ è punto di sella, vale a dire che è stazionario ma non è né punto di massimo relativo, né punto di minimo relativo.

Osserviamo che il viceversa delle implicazioni (i) $\to$ (v) è sempre falso. Esempi (in cui $\underline{x}_0 = \underline{0}$ e λ_1, λ_2 sono gli autovalori di $H_f(\underline{0})$):

(i) $f(x,y)=x^4-y^2$; $\lambda_1=0$, $\lambda_2=-2$ ma $f(0,0)<f(\frac{1}{n},0)$ $\forall n\in\mathbb{N}^+$;

(ii) $f(x,y)=y^2-x^4$; $\lambda_1=0$, $\lambda_2=2$ ma $f(0,0)>f(\frac{1}{n},0)$ $\forall n\in\mathbb{N}^+$;

(iii) $f(x,y)=x^4+y^4$; $f(0,0)\le f(x,y)$ $\forall(x,y)\in\mathbb{R}^2$, ma $\lambda_1=\lambda_2=0$;

(iv) $f(x,y)=-x^4-y^4$; $f(0,0)\ge f(x,y)$ $\forall(x,y)\in\mathbb{R}^2$, ma $\lambda_1=\lambda_2=0$;

(v) $f(x,y)=-x^4+y^4$; $f(0,\frac{1}{n})>f(0,0)>f(\frac{1}{n},0)$ $\forall n\in\mathbb{N}^+$, ma $\lambda_1=\lambda_2=0$.

Perché il teorema funziona?

Sia $f\in C^2(A)$. Per queste funzioni si ha una formula (formula di Taylor) che è una specie di differenziabilità del 2° ordine:

$$f(\underline{x})-f(\underline{x}_0)=\langle\nabla f(\underline{x}_0),\underline{x}-\underline{x}_0\rangle_N+\frac{1}{2}\langle H_f(\underline{x}_0)\cdot(\underline{x}-\underline{x}_0),\underline{x}-\underline{x}_0\rangle_N+\omega_2(\underline{x}),$$

ove

$$\lim_{\underline{x}\to\underline{x}_0}\frac{\omega(\underline{x})}{|\underline{x}-\underline{x}_0|_N^2}=0.$$

Dunque se $f\in C^2(A)$, nell'intorno di ogni $\underline{x}_0\in A$ si può approssimare f con un polinomio di 2° grado, e la differenza è infinitesima per $\underline{x}\to\underline{x}_0$ di ordine superiore a $|\underline{x}-\underline{x}_0|_N^2$.

Allora, se $\underline{x}_0$ è punto stazionario per f, si ha $\nabla f(\underline{x}_0)=0$ e dunque

$$f(\underline{x})-f(\underline{x}_0)=\frac{1}{2}\langle H_f(\underline{x}_0)(\underline{x}-\underline{x}_0),\underline{x}-\underline{x}_0\rangle_N+\omega_2(\underline{x});$$

dunque per valori di $\underline{x}$ vicini a $\underline{x}_0$ il segno di $f(\underline{x})-f(\underline{x}_0)$ è determinato dal termine quadratico $\frac{1}{2}\langle H_f(\underline{x}_0)(\underline{x}-\underline{x}_0),\underline{x}-\underline{x}_0\rangle_N$,

perchè il contributo di $\omega_2(\underline{x})$ è trascurabile.
Perciò: se gli autovalori di $H_f(\underline{x}_0)$ sono tutti positivi, allora

$$\langle H_f(\underline{x}_0)(\underline{x}-\underline{x}_0), \underline{x}-\underline{x}_0\rangle_N > 0 \quad \forall\, \underline{x} \in \mathbb{R}^N \setminus \{\underline{x}_0\},$$

per cui

$$f(\underline{x}) > f(\underline{x}_0) \quad \forall\, \underline{x} \in B(\underline{x}_0, \delta) \setminus \{\underline{x}_0\},$$

purchè $\delta > 0$ sia abbastanza piccolo.
Se gli autovalori di $H_f(\underline{x}_0)$ sono tutti negativi, allora

$$\langle H_f(\underline{x}_0)(\underline{x}-\underline{x}_0), \underline{x}-\underline{x}_0\rangle_N < 0 \quad \forall\, \underline{x} \in \mathbb{R}^N \setminus \{\underline{x}_0\},$$

per cui

$$f(\underline{x}) < f(\underline{x}_0) \quad \forall\, \underline{x} \in B(\underline{x}_0, \delta) \setminus \{\underline{x}_0\},$$

purchè $\delta > 0$ sia abbastanza piccolo.
Se infine $H_f(\underline{x}_0)$ ha sia autovalori positivi, sia autovalori negativi, allora $\langle H_f(\underline{x}_0)(\underline{x}-\underline{x}_0), \underline{x}-\underline{x}_0\rangle_N$ cambia segno in ogni palla $B(\underline{x}_0, \delta)$, e il punto $\underline{x}_0$ è di sella.

Dunque è importante saper decidere il segno degli autovalori della matrice simmetrica $H_f(\underline{x}_0)$.

<u>Caso $N=2$</u> Poniamo $A = H_f(\underline{x}_0) = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$. Gli autovalori di A verificano

$$\det(A-\lambda I) = \det\begin{pmatrix} a-\lambda & b \\ b & c-\lambda \end{pmatrix} = (a-\lambda)(c-\lambda) - b^2 = \lambda^2 - (a+c)\lambda + ac - b^2 = 0.$$

(I) Se $\det A > 0$ e $a > 0$, gli autovalori di A sono positivi

Infatti $\det A = ac - b^2 > 0$, quindi $ac > b^2 \geq 0$, da cui (essendo $a > 0$) segue $c > 0$. Perciò $a + c > 0$; dunque, se λ è autovalore, ossia $\lambda^2 - (a+c)\lambda + ac - b^2 = 0$, si ha

$$\lambda = 0 \Rightarrow ac - b^2 = 0, \text{ assurdo},$$

$$\lambda < 0 \Rightarrow \lambda^2 = (a+c)\lambda - (ac - b^2) < 0, \text{ assurdo},$$

ossia $\lambda > 0$.

(II) Se $\det A > 0$ e $a < 0$, gli autovalori di A sono negativi.

Infatti $\det A = ac - b^2 > 0$, quindi $ac > b^2 \geq 0$, da cui (essendo $a < 0$) segue $c < 0$. Perciò $a + c < 0$; dunque, se λ è autovalore, ossia $\lambda^2 - (a+c)\lambda + ac - b^2 = 0$, si ha

$$\lambda = 0 \Rightarrow ac - b^2 = 0, \text{ assurdo};$$

$$\lambda > 0 \Rightarrow \lambda^2 = (a+c)\lambda - (ac - b^2) < 0, \text{ assurdo},$$

ossia $\lambda < 0$.

(III) Se $\det A < 0$, allora gli autovalori di A sono uno positivo e uno negativo.

Infatti, per ogni matrice quadrata $N \times N$, si dimostra che $\det A = \prod_{j=1}^{N} \lambda_j$, ove λ_j sono gli autovalori di A; quindi $0 > \det A = \lambda_1 \lambda_2$, e pertanto i due autovalori hanno segno discorde.

(IV) Se $\det H = 0$, almeno un autovalore è nullo, e non possiamo dire nulla sulla natura del punto stazionario, come mostrano gli esempi (iii)-(iv)-(v) di pagina 31.
L'unica cosa che si può dire quando $\det H = 0$ è:

- se un autovalore è nullo e l'altro è positivo, allora $\underline{x_0}$ <u>non è</u> punto di massimo relativo per f;
- se un autovalore è nullo e l'altro è negativo, allora $\underline{x_0}$ <u>non è</u> punto di minimo relativo per f.

<u>Esercizi</u>

1. Determinare i punti di massimo e minimo relativo per

$$f(x,y) = 2x^2 - 3xy + ay^2 - 4x + y,$$

al variare del parametro $a \in \mathbb{R}$.

2. Determinare i punti di massimo e minimo relativo per

(i) $f_1(x,y) = x^2 + 2y^2 - 4x - 4y$

(ii) $f_2(x,y) = x^2 y\, e^{-x^2-y^2}$

(iii) $f_3(x,y,z) = -x - y - z + xyz$

3. Determinare le dimensioni di una scatola rettangolare di volume V fissato e area (della base più le 4 facce laterali) minima.

4. Trovare il rettangolo con i lati paralleli agli assi cartesiani, inscritto nell'ellisse di equazione $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$, di area massima.

<u>Caso $N=3$</u> Sia $A = H_f(x_0) = \begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix}$

Consideriamo in sequenza i determinanti delle 3 matrici nei riquadri: a, $ad-b^2$, $\det A$.

(i) Essi devono essere tutti e 3 diversi da 0.

(ii) Si mettono in fila un segno $+$, seguito dai segni dei 3 numeri.

(iii) Si conta il numero N^+ delle <u>permanenze</u> e il numero N^- delle <u>variazioni</u> di segno: chiaramente $N^+ + N^- = 3$ e $N^+ \geq 0$, $N^- \geq 0$.

Allora la matrice A ha N^+ autovalori positivi e N^- autovalori negativi.

<u>Osservazione</u> se $\det H = 0$, allora un autovalore (almeno) è nullo. Se gli altri due sono diversi da 0 e discordi, allora il punto è di sella. Se gli altri due sono concordi, non si può dire nulla.

Esempio Sia $f(x,y,z) = \frac{x^2}{2} - xy + 2y^2 + 2yz + k\frac{z^2}{2}$, con $k \in \mathbb{R}$ fissato. Si ha

$$\nabla f(x,y,z) = \left(x-y,\ -x+4y+2z,\ 2y+kz\right),$$

e i punti stazionari sono le soluzioni del sistema

$$\begin{cases} x-y=0 \\ -x+4y+2z=0 \\ 2y+kz=0, \end{cases}$$

vale a dire, se $k \neq \frac{4}{3}$, $(x,y,z) = (0,0,0)$. Inoltre

$$H_f(x,y,z) = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 4 & 2 \\ 0 & 2 & k \end{pmatrix} \quad \text{per ogni } (x,y,z) \in \mathbb{R}^3.$$

I tre determinanti sono

$$a = 1, \qquad ad - b^2 = 3, \qquad \det A = 3k - 4.$$

Fissando ad esempio $k=2$, la sequenza dei segni è $++++$; quindi gli autovalori sono tutti e tre positivi, e $(0,0,0)$ è punto di minimo relativo per f. Scegliendo invece $a = 0$, la sequenza diventa $+++-$, per cui un autovalore diventa negativo e $(0,0,0)$ è ora punto di sella.

Esercizio Trovare i punti stazionari, stabilendone la natura, per le funzioni

(i) $f(x,y) = x^4 + y^4 - 2(x+y)^2 + 2$,

(ii) $f(x,y,z) = x^2 - 2x + y^2 + \ln(1+z^2)$.

Risoluzione

(i) La f è di classe C^∞ e i suoi punti stazionari verificano

$$\begin{cases} f_x(x,y) = 4x^3 - 4(x+y) = 0 \\ f_y(x,y) = 4y^3 - 4(x+y) = 0. \end{cases}$$

Dunque $4(x+y) = 4x^3 = 4y^3$, da cui $x=y$ e $4x^3 = 4(2x)$, ossia $x=0$ oppure $x = \pm\sqrt{2}$. Si hanno perciò i tre punti stazionari

$$(0,0), \quad (\sqrt{2},\sqrt{2}), \quad (-\sqrt{2},-\sqrt{2}).$$

Calcoliamo la matrice Hessiana:

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

da cui

$$H_f(0,0) = \begin{pmatrix} -4 & -4 \\ -4 & -4 \end{pmatrix}, \quad H_f(\sqrt{2},\sqrt{2}) = \begin{pmatrix} 20 & -4 \\ -4 & 20 \end{pmatrix} = H_f(-\sqrt{2},-\sqrt{2}).$$

Nei punti $\pm(\sqrt{2},\sqrt{2})$ si ha $\det H_f > 0$, $a=20$, e dunque tali punti sono di minimo relativo. Nel punto $(0,0)$ si ha $\det H_f = 0$, e ciò non ci dà informazioni. Tuttavia per $\varepsilon > 0$ (piccolo) si ha

$f(0,\varepsilon) = \varepsilon^4 - 2\varepsilon^2 + 2 = 2 - \varepsilon^2(2-\varepsilon^2) < 2,$

$f(0,0) = 2$

$f(\varepsilon,\varepsilon) = 2\varepsilon^4 + 2 > 2,$

e quindi $(0,0)$ è punto di sella.

(ii) La f è di classe C^∞. I punti stazionari risolvono

$$\nabla f(x,y,z) = \underline{0} \iff \begin{cases} 2x-2=0 \\ 2y=0 \\ \dfrac{2z}{1+z^2}=0 \end{cases} \iff \begin{cases} x=1 \\ y=0 \\ z=0 \end{cases}.$$

Il punto $(1,0,0)$ è stazionario. Dato che

$$H_f(x,y,z) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & \dfrac{2-2z^2}{(1+z^2)^2} \end{pmatrix}, \quad H_f(0,0,0) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix},$$

il punto $(1,0,0)$ è di minimo relativo (perché gli autovalori sono $\lambda_1=\lambda_2=\lambda_3=2$). Ma $(1,0,0)$ è anche di minimo assoluto, essendo ...

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

$$f(x,y,z) = (x-1)^2 - 1 + y^2 + \ln(1+z^2) \geq -1 = f(1,0,0)$$

per ogni $(x,y,z) \in \mathbb{R}^3$.