Massimi e minimi

Siano dati l'insieme $X \subseteq I\!R^n$ e la funzione $f\colon X \to I\!R$.

Definizione 1. $\overline{x}^0 \in X$ *si dice*

1) *punto di* massimo globale o assoluto *se e solo se* $f(\overline{x}) \leq f(\overline{x}^0)$ *per ogni* $\overline{x} \in X$*;*
2) *punto di* massimo locale o relativo *se e solo se esiste in intorno* $B(\overline{x}^0, r)$ *tale che* $f(\overline{x}) \leq f(\overline{x}^0)$ *per ogni* $\overline{x} \in X \cap B(\overline{x}^0, r)$*.*

In modo analogo si definiscono i punti di *minimo globale* e *locale.*

Osservazioni.

1) Questi punti vengono genericamente chiamati *estremi* (globali e locali).
2) Se $\overline{x}^0$ è un punto di estremo globale allora è anche un punto di estremo locale ma non vale il viceversa.
3) Se X è compatto e f è continua allora (grazie al Teorema di Weierstrass) esistono gli estremi assoluti.

Esempi.

1) L'origine è un punto di minimo globale per la funzione $f(x, y) = \sqrt{x^2 + y^2}$.
2) Ogni punto di coordinate $(0, y)$, $y \in I\!R$ è un punto di massimo globale per la funzione $f(x, y) = -x^2$.

Definizione 2. *Sia* $A \subseteq I\!R^n$ *aperto e* $f\colon A \to I\!R$ *differenziabile in* $\overline{x}^0 \in A$*. Il punto* $\overline{x}^0$ *si dice* punto stazionario *se* $df(\overline{x}^0) = 0$.

Osservazioni.

1) $df(\overline{x}^0) = 0$ se e solo se $\nabla f(\overline{x}^0) = 0$.
2) Analogamente al caso unidimensionale, $\overline{x}^0$ è stazionario se e solo se l'iperpiano tangente al grafico di f in $(\overline{x}^0, f(\overline{x}^0))$ è ortogonale all'ultimo elemento della base canonica di $I\!R^{n+1}$ (per $n = 2$ ciò significa che il piano tangente è orizzontale).

Teorema 1. *Sia* $A \subseteq I\!R^n$ *aperto e* $f\colon A \to I\!R$ *differenziabile in* $\overline{x}^0 \in A$*. Se* $\overline{x}^0$ *è un punto di estremo locale allora è anche un punto stazionario.*

Dimostrazione. Sia $\overline{x}^0$ un estremo locale per f. Si fissi un versore arbitrario $\overline{v} \in I\!R^n$. Allora la funzione $t \mapsto f(\overline{x}^0 + t\overline{v})$ è definita in un intorno dello zero, inoltre lo zero è un punto di estremo locale; per cui $\frac{d}{dt} f(\overline{x}^0 + t\overline{v})_{|t=0} = 0$. Quindi

$$df(\overline{x}^0)(\overline{v}) = D_{\overline{v}}(\overline{x}) = \frac{d}{dt} f(\overline{x}^0 + t\overline{v})_{|t=0} = 0.$$

Grazie all'arbitrarietà del versore $\overline{v}$ segue la tesi. ∎

Osservazione. Un punto può essere stazionario e non essere un punto di estremo locale; ad esempio $x = 0$ per $f(x) = x^3$ (punto di flesso a tangente orizzontale) oppure $(x, y) = (0, 0)$ per $f(x, y) = y^2 - x^2$.

Definizione 3. *Un punto* $\overline{x}^0$ *si dice* di sella *o* di colle *se è stazionario ma non è un estremo locale.*

Osservazione. Un punto $\overline{x}^0$ stazionario per f è di sella se e solo se in ogni suo intorno (contenuto nel dominio della funzione) f assume sia valori maggiori che minori di $f(\overline{x}^0)$.

Se f è differenziabile almeno due volte le sue derivate seconde possono fornire informazioni sulla natura dei punti stazionari. Prima di enunciare il cosiddetto "test delle derivate seconde" riprendiamo brevemente un concetto legato alle matrici simmetriche.

Definizione 4. *Una matrice* $n \times n$ *simmetrica* A *si dice:*
1) definita positiva *se* $\overline{h}^T A \, \overline{h} > 0$ *per ogni* $\overline{h} \in I\!R^n$ *e* $\overline{h} \neq 0$*;*
2) semidefinita positiva *se* $\overline{h}^T A \, \overline{h} \geq 0$ *per ogni* $\overline{h} \in I\!R^n$*;*
3) definita negativa *se* $\overline{h}^T A \, \overline{h} < 0$ *per ogni* $\overline{h} \in I\!R^n$ *e* $\overline{h} \neq 0$*;*
4) semidefinita negativa *se* $\overline{h}^T A \, \overline{h} \leq 0$ *per ogni* $\overline{h} \in I\!R^n$*;*
5) indefinita *se esistono* $\overline{h}_1, \overline{h}_2 \in I\!R^n$ *tali che* $\overline{h}_1^T A \, \overline{h}_1 > 0$ *e* $\overline{h}_2^T A \, \overline{h}_2 < 0$.

Osservazione. Si noti che se A è definita allora è anche semidefinita (ma non vale il viceversa).

Queste proprietà possono essere espresse in termini degli autovalori di A.

Proposizione 1. *Data una matrice* $n \times n$ *simmetrica* A*, allora:*
1) A *è definita positiva se e solo se i suoi autovalori sono* positivi*;*
2) A *è semidefinita positiva se e solo se i suoi autovalori sono* positivi *o* nulli*;*
3) A *è definita negativa se e solo se i suoi autovalori sono* negativi*;*
4) A *è semidefinita negativa se e solo se i suoi autovalori sono* negativi *o* nulli*;*
5) A *è indefinita se e solo se esistono due autovalori di* segno opposto.

Esempio. La matrice $\begin{pmatrix} 1 & 1 & 2 \\ 0 & 5 & -1 \\ 0 & 0 & 7 \end{pmatrix}$ è definita positiva.

Nel caso $n = 2$ sarà utile il seguente criterio.

Proposizione 2. *Sia* $A = \left(\begin{smallmatrix} a & b \\ b & c \end{smallmatrix}\right)$*, allora:*
1) se $\det A > 0$ *e* $a > 0$ *allora* A *è definita positiva;*
2) se $\det A > 0$ *e* $a < 0$ *allora* A *è definita negativa;*
3) se $\det A < 0$ *allora* A *è indefinita;*
5) se $\det A = 0$ *allora* A *è semidefinita (positiva o negativa) ma non definita;*

Proposizione 3. *Sia* A *una matrice* $n \times n$ *simmetrica e siano* λ_{min} *e* λ_{max} *rispettivamente il più piccolo e il più grande dei suoi autovalori. Allora, per ogni* $\overline{h} \in I\!R^n$ *vale*

$$\lambda_{min} \|\overline{h}\|^2 \leq \overline{h}^T A \, \overline{h} \leq \lambda_{max} \|\overline{h}\|^2.$$

Dimostrazione. Siano $\lambda_1, \ldots, \lambda_n$ gli autovalori di A (ripetuti in base alla loro molteplicità) e sia Q una matrice ortogonale che diagonalizza A. Allora $Q^T A \, Q = \Lambda$, dove Λ è una matrice diagonale la cui diagonale principale è costituita dagli autovalori di A. Sia $\overline{h} \in I\!R^n$ arbitrario e poniamo $(k_1, \ldots, k_n) = \overline{k} = Q^T \overline{h}$, si noti che $\|\overline{k}\| = \|\overline{h}\|$. Abbiamo

$$\overline{h}^T A\, \overline{h} = \overline{h}^T Q Q^T A Q Q^T \overline{h} = (Q^T \overline{h})^T Q^T A Q (Q^T \overline{h}) = \overline{k}^T \Lambda\, \overline{k} = \lambda_1 k_1^2 + \cdots + \lambda_n k_n^2.$$

D'altra parte

$$\lambda_{min} \|\overline{h}\|^2 = \lambda_{min} \|\overline{k}\|^2 \leq \lambda_1 k_1^2 + \cdots + \lambda_n k_n^2 \leq \lambda_{max} \|\overline{k}\|^2 = \lambda_{max} \|\overline{h}\|^2,$$

che conclude la dimostrazione. ∎

Osservazione. Nel seguito useremo la tesi della proposizione precedente nella forma

$$\lambda_{min} \leq \left(\frac{\overline{h}}{\|\overline{h}\|} \right)^T A \left(\frac{\overline{h}}{\|\overline{h}\|} \right) \leq \lambda_{max} \qquad \forall\, \overline{h} \in I\!R^n,\ \overline{h} \neq 0.$$

Ricordiamo che il differenziale secondo $d^2 f(\overline{x}^0)$ è una forma quadratica su $I\!R^n$ rappresentato, in coordinate, dalla matrice Hessiana $H_f(\overline{x}^0)$.

Teorema 2. *Sia $A \subseteq I\!R^n$ aperto, $f \colon A \to I\!R$ una funzione di classe C^2 e $\overline{x}^0$ un punto stazionario per f. Valgono le seguenti implicazioni:*
1) se $d^2 f(\overline{x}^0)$ è definita positiva allora $\overline{x}^0$ è un punto di minimo locale stretto;
2) se $d^2 f(\overline{x}^0)$ è definita negativa allora $\overline{x}^0$ è un punto di massimo locale stretto;
3) se $d^2 f(\overline{x}^0)$ è indefinita allora $\overline{x}^0$ è un punto di sella.

Dimostrazione. Caso 1), sia $d^2 f(\overline{x}^0)$ definita positiva, allora gli autovalori della matrice Hessiana $H_f(\overline{x}^0)$ sono tutti positivi. Sia λ_{min} il più piccolo. Avremo

$$\left(\frac{\overline{h}}{\|\overline{h}\|} \right)^T H_f(\overline{x}^0) \left(\frac{\overline{h}}{\|\overline{h}\|} \right) \geq \lambda_{min} > 0 \qquad \forall\, \overline{h} \in I\!R^n,\ \overline{h} \neq 0,$$

quindi, prendendo lo sviluppo di Taylor di ordine 2 di f attorno al punto $\overline{x}^0$, troviamo

$$\begin{aligned} f(\overline{x}^0 + \overline{h}) - f(\overline{x}^0) &= \frac{1}{2} \overline{h}^T H_f(\overline{x}^0) \overline{h} + o(\|\overline{h}\|^2) \\ &= \|\overline{h}\|^2 \left(\frac{1}{2} \left(\frac{\overline{h}}{\|\overline{h}\|} \right)^T H_f(\overline{x}^0) \left(\frac{\overline{h}}{\|\overline{h}\|} \right) + o(1) \right) \\ &\geq \|\overline{h}\|^2 \left(\frac{\lambda_{min}}{2} + o(1) \right) \quad \text{per } \overline{h} \to 0. \end{aligned}$$

Di conseguenza, per $\|\overline{h}\|$ abbastanza piccolo $f(\overline{x}^0 + \overline{h}) - f(\overline{x}^0) > 0$, cioè $\overline{x}^0$ è un punto di minimo locale stretto. Nel caso 2) si ragiona in modo completamente analogo.
Caso 3), sia $d^2 f(\overline{x}^0)$ indefinita. Esisteranno allora due versori $\overline{h}_1$ e $\overline{h}_2$ tali che $\overline{h}_1^T H_f(\overline{x}^0) \overline{h}_1 > 0$

e $\overline{h}_2^T H_f(\overline{x}^0)\overline{h}_2 < 0$. Sfruttando nuovamente lo sviluppo di Taylor di ordine 2 di f attorno a $\overline{x}^0$ abbiamo, per $|t|$ abbastanza piccolo,

$$f(\overline{x}^0 + t\overline{h}_1) - f(\overline{x}^0) = t^2 \left(\frac{1}{2}\overline{h}_1^T H_f(\overline{x}^0)\overline{h}_1 + o(1) \right) > 0.$$

Quindi la funzione f ristretta ad un opportuno segmento centrato in $\overline{x}^0$ e di direzione $\overline{h}_1$ ha in $\overline{x}^0$ un minimo stretto. Allora esistono punti arbitrariamente vicini ad $\overline{x}^0$ in cui f assume valori maggiori di $f(\overline{x}^0)$. In modo simile, scambiando $\overline{h}_1$ con $\overline{h}_2$ si deduce che esistono punti arbitrariamente vicini ad $\overline{x}^0$ in cui f assume valori minori di $f(\overline{x}^0)$, dunque $\overline{x}^0$ è un punto di sella.
■

Osservazioni.

1) Nel caso in cui $d^2 f(\overline{x}^0)$ è semidefinita ma non definita e non è nulla, è ancora possibile ricavare un'informazione parziale, infatti si può dimostrare che:
   1) se $d^2 f(\overline{x}^0)$ è semidefinita positiva e non è nulla allora $\overline{x}^0$ è un punto di minimo locale o di sella;
   2) se $d^2 f(\overline{x}^0)$ è semidefinita negativa e non è nulla allora $\overline{x}^0$ è un punto di massimo locale o di sella.
   Ad esempio, le funzioni $f(x,y) = x^2$, $f(x,y) = y^2 - x^4$ in $(0,0)$ rientrano nel caso 1) e hanno, rispettivamente, un minimo locale e un punto di sella.
2) Quando $d^2 f(\overline{x}^0) = 0$ non si ha nessuna informazione. Infatti le funzioni $f(x,y) = x^4 + y^4$, $f(x,y) = -(x^4 + y^4)$ e $f(x,y) = x^4 - y^4$ hanno in $(0,0)$ il differenziale secondo nullo e, rispettivamente, un punto di minimo, un punto di massimo e un punto di sella.

Ricordando la Proposizione 2 ricaviamo il seguente corollario.

Corollario 1. *Sia $f(x,y)$ una funzione di classe C^2 definita in un aperto A di $\mathbb{R}^2$ e (x_0,y_0) un punto stazionario. Allora:*

1) se $\det H_f(x_0,y_0) > 0$ e $f_{xx}(x_0,y_0) > 0$ allora (x_0,y_0) è un punto di minimo locale stretto;

2) se $\det H_f(x_0,y_0) > 0$ e $f_{xx}(x_0,y_0) < 0$ allora (x_0,y_0) è un punto di massimo locale stretto;

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

Esempio. La funzione $f(x,y) = 2x^3 - 6xy + 3y^2$ ha i punti stazionari $(0,0)$ e $(1,1)$. Abbiamo $H_f(0,0) = \begin{pmatrix} 0 & -6 \\ -6 & 6 \end{pmatrix}$ e $H_f(1,1) = \begin{pmatrix} 12 & -6 \\ -6 & 6 \end{pmatrix}$. Grazie alla proposizione precedente si deduce che $(0,0)$ è un punto di sella e che $(1,1)$ è un punto di minimo locale.