UNIVERSITÀ DEGLI STUDI DI FERRARA
C.d.S. Ingegneria Civile e Ambientale

Appunti del corso di Analisi Matematica II [1]

Michele Miranda
Dipartimento di Matematica e Informatica
via Machiavelli 30, I-44121 Ferrara
e–mail: michele.miranda@unife.it

a.a. 2019-2020

[1]Versione aggiornata al 29 ottobre 2019

Nel presente fascicolo sono raccolti gli appunti del corso di Analisi Matematica 2 dal Corso di Laurea Triennale di Ingegneria Civile e Ambientale dell'Università di Ferrara.

Il materiale contenuto in queste note vuole essere semplicemente una guida relativa agli argomenti trattati durante il corso; è inevitabilmente incompleto, così come è inevitabile che siano presenti errori ed inesattezze. Non si risponde tuttavia degli errori che possono essere contenuti in questo fascicolo, in quanto è cura del lettore rilevare e segnalare eventuali imprecisioni.

I presenti appunti non hanno la pretesa di sostituire un buon libro di testo, che resta indispensabile per acquisire una conoscenza dignitosa della materia. Viene fornita in bibliografia una lista di testi che si ritengono validi per lo studio della materia. La funzione di questi appunti è piuttosto quella di facilitare gli studenti e indicare loro il bagaglio *minimo* di conoscenze richieste per affrontare l'esame. Si consiglia pertanto sempre di studiare sui testi di Analisi Matematica esistenti in letteratura, sicuramente più affidabili e corretti; fortunatamente le biblioteche dei nostri Atenei sono molto buone e ben fornite di ottimi testi. Tra i vari testi disponibili consigliamo sicuramente il testo [2] o la sua edizione precedente [1]; un altro ottimo testo è [3]

Si consiglia infine di prestare attenzione alla data di aggiornamento della presente dispensa, in quanto è possibile che alcune parti vengano corrette ed integrate durante il corso.

Michele Miranda
Ferrara, 29 ottobre 2019

Indice

Capitolo 5

Estremi e punti stazionari

In questo capitolo ci occuperemo di problemi di massimo e minimo per funzioni di più variabili; in particolare, vedremo come determinare tali punti, detti anche estremi (assoluti e locali), mediante l'individuazione dei punti stazionari della funzione, sia liberi che vincolati. Vedremo come classificare i punti stazionari liberi mediante l'uso della matrice Hessiana e come determinare i punti stazionari vincolati introducendo i metodi detti per parametrizzazione del vincolo e dei moltiplicatori di Lagrange.

5.1 Massimi e minimi

Richiamiamo anzitutto alcune definizioni; data una funzione $f : E \to \mathbb{R}$, con $E \subset \mathbb{R}^n$ e $x_0 \in E$. Diremo che

1. x_0 è un *punto di minimo assoluto o globale* su E (rispettivamente *punto di massimo assoluto o globale* su E), o semplicemente *minimo* (risp. *massimo*) su E,

$$f(x_0) \leq f(x) \text{ (risp. } f(x_0) \geq f(x)), \qquad \forall x \in E;$$

2. x_0 è un *punto di minimo assoluto o globale stretto* su E (rispettivamente *punto di massimo assoluto o globale stretto* su E), o semplicemente *minimo stretto* (risp. *massimo stretto*) su E,

$$f(x_0) < f(x) \text{ (risp. } f(x_0) > f(x)), \qquad \forall x \in E \setminus \{x_0\};$$

3. $x_0 \in E$ è *punto di minimo locale* (risp. *massimo locale*) se esiste $r > 0$ tale che

$$f(x_0) \leq f(x) \text{ (risp. } f(x_0) \geq f(x)), \qquad \forall x \in E \cap B_r(x_0);$$

4. $x_0 \in E$ è *punto di minimo locale stretto* (risp. *massimo locale stretto*) se esiste $r > 0$ tale che

$$f(x_0) < f(x) \text{ (risp. } f(x_0) > f(x)), \qquad \forall x \in E \cap B_r(x_0) \setminus \{x_0\}.$$

Un punto di massimo o minimo (assoluto o locale) viene anche detto estremo (assoluto o locale). Infine, nel caso di di x_0 punto estremo assoluto, il valore $f(x_0)$ si dirà rispettivamente *valore minimo* o *valore massimo* e scriveremo rispettivamente

$$f(x_0) = \min_E f, \qquad f(x_0) = \max_E f.$$

Si dirà poi che f ammette minimo su E (risp. ammette massimo) se esiste almeno un punto di minimo (risp. massimo) assoluto su E.

Per quanto riguarda l'esistenza di massimo e minimo abbiamo il seguente risultato.

Teorema 5.1 (Weierstrass) *Sia $f : E \to \mathbb{R}$ una funzione continua con $E \subset \mathbb{R}^n$ compatto; allora f ammette minimo e massimo su E.*

Dimostrazione. Dimostriamo l'esistenza di un punto di minimo; la dimostrazione dell'esistenza di un punto di massimo sarà analoga.

Per semplificare la notazione supponiamo che $n = 2$ e definiamo

$$\alpha = \inf_E f = \inf\{f(x) : x \in E\};$$

dobbiamo dimostrare che esiste $(x_0, y_0) \in E$ per cui $f(x_0, y_0) = \alpha$. La compattezza di E implica la sua limitatezza, e quindi possiamo racchiudere E in un rettangolo $R = [a, b] \times [c, d]$, cioè

$$E \subset R.$$

Dividiamo R in 4 rettangoli uguali,

$$R = Q_1 \cup Q_2 \cup Q_3 \cup Q_4$$

con

$$Q_1 = \left[a, a + \frac{b-a}{2}\right] \times \left[c, c + \frac{d-c}{2}\right], \quad Q_2 = \left[a + \frac{b-a}{2}, b\right] \times \left[c, c + \frac{d-c}{2}\right]$$
$$Q_3 = \left[a, a + \frac{b-a}{2}\right] \times \left[c + \frac{d-c}{2}, d\right], \quad Q_4 = \left[a + \frac{b-a}{2}, b\right] \times \left[c + \frac{d-c}{2}, d\right].$$

Definiamo anche $C_i = E \cap Q_i$, $i = 1, 2, 3, 4$. Dato che $E = C_1 \cup C_2 \cup C_3 \cup C_4$, avremo che per uno dei quattro indici $i = 1, 2, 3, 4$ deve valere che $C_i \neq \emptyset$ e

$$\inf_E f = \inf_{C_i} f.$$

Poniamo quindi con tale scelta dell'indice i

$$E_1 = C_i, \quad R_1 = Q_i := [a_1, b_1] \times [c_1, d_1].$$

Con tale definizione abbiamo che

$$a \le a_1, b \ge b_1, c \le c_1, d \ge d_1$$
$$b_1 - a_1 = \frac{b-a}{2}, \qquad d_1 - c_1 = \frac{d-c}{2}.$$

Possiamo iterare questo ragionamento dividendo R_1 in quattro rettangoli uguali in modo da definire un insieme non vuoto $E_2 = E_1 \cap R_2 = E \cap R_2$ con $R_2 = [a_2, b_2] \times [c_2, d_2]$ tali che

$$\inf_E f = \inf_{E_2} f$$

e

$$a_1 \le a_2, b_1 \ge b_2, c_1 \le c_2, d_1 \ge d_2$$
$$b_2 - a_2 = \frac{b_1 - a_1}{2} = \frac{b-a}{4}, \qquad d_2 - c_2 = \frac{d_1 - c_1}{2} = \frac{d-c}{4}.$$

Iterando ulteriormente la procedura si costruisce una successione di insiemi non vuoti $E_h = E_{h-1} \cap R_h = E \cap R_h$ con $R_h = [a_h, b_h] \times [c_h, d_h]$ tali che

$$\inf_E f = \inf_{E_h} f$$

e

$$a_{h-1} \le a_h, b_{h-1} \ge b_h, c_{h-1} \le c_h, d_{h-1} \ge d_h$$

$$b_h - a_h = \frac{b_{h-1} - a_{h-1}}{2} = \frac{b-a}{2^h}, \qquad d_h - c_h = \frac{d_{h-1} - c_{h-1}}{2} = \frac{d-c}{2^h}.$$

Le successioni $(a_h)_{h\in\mathbb{N}}$ e $(c_h)_{h\in\mathbb{N}}$ sono monotone non decrescenti, mentre le successsioni $(b_h)_{h\in\mathbb{N}}$ e $(d_h)_{h\in\mathbb{N}}$ sono monotone non crescenti. Quindi tali successioni sono regolari e dato che $a_h \leq b_h$, $c_h \leq d_h$, convergono a dei numeri reali. Il fatto che

$$\lim_{h\to+\infty}(b_h - a_h) = \lim_{h\to+\infty}\frac{b-a}{2^h} = 0$$

implica che le due successioni $(a_h)_{h\in\mathbb{N}}$ e $(b_h)_{h\in\mathbb{N}}$ hanno lo stesso limite che chiameremo x_0, mentre il fatto che

$$\lim_{h\to+\infty}(d_h - c_h) = \lim_{h\to+\infty}\frac{d-c}{2^h} = 0$$

implica che le due successioni $(c_h)_{h\in\mathbb{N}}$ e $(d_h)_{h\in\mathbb{N}}$ hanno lo stesso limite che chiameremo y_0. Mostriamo anzitutto che $(x_0, y_0) \in E$; siccome gli insiemi E_h sono non vuoti, posso prendere un punto $(x_h, y_h) \in E_h$ per ogni $h \in \mathbb{N}$. Le disuguaglianze

$$a_h \leq x_h \leq b_h, \qquad c_h \leq y_h \leq d_h$$

implicano che $x_h \to x_0$ e $y_h \to y_0$, cioè $(x_h, y_h) \to (x_0, y_0)$ che quindi appartiene ad E perchè E è chiuso. Si noti che $(x_0, y_0) \in R_h$ per ogni $h \in \mathbb{N}$ e che il fatto che $(x_0 m y_0) \in E$ implica che

$$\alpha \leq f(x_0, y_0).$$

Siccome f è continua in particolare in (x_0, y_0) sappiamo che per ogni $\varepsilon > 0$ esiste $r > 0$ tale che

$$|f(x,y) - f(x_0,y_0)| < \varepsilon \qquad \forall (x,y) \in Q_r(x_0,y_0).$$

In particolare su $Q_r(x_0, y_0)$ vale la disuguaglianza

$$f(x_0, y_0) - \varepsilon < f(x,y). \tag{5.1}$$

Osserviamo infine che se h_0 è scelto in modo tale che

$$\frac{b-a}{2^{h_0}} < r, \qquad \frac{d-c}{2^{h_0}} < r,$$

allora per ogni $h \geq h_0$ si ha che $R_h \subset Q_r(x_0, y_0)$, e quindi grazie a (5.1)

$$f(x_0, y_0) - \varepsilon < \inf_{E_h} f = \inf_E f = \alpha.$$

In definitiva abbiamo dimostrato che per ogni $\varepsilon > 0$

$$\alpha \leq f(x_0, y_0) < \alpha + \varepsilon,$$

e quindi $f(x_0, y_0) = \alpha$ che dimostra che (x_0, y_0) è un punto di minimo assoluto o globale su E. □

Esistono versioni del Teorema di Weierstrass che garantiscono l'esistenza di massimi e minimi anche senza la richiesta di compattezza; presentiamo qui la seguente versione modificata.

Proposizione 5.2 *Sia $f : E \to \mathbb{R}$ una funzione continua. Se esistono un compatto $K \subset E$ e $x_1 \in K$ tali che $f(x) \geq f(x_1)$ per ogni $x \in E \setminus K$, allora esiste il minimo di f su E. Analogamomante, se esistono un compatto $K \subset E$ e $x_1 \in K$ tali che $f(x) \leq f(x_1)$ per ogni $x \in E \setminus K$ allora esiste il massimo di f su E.*

Dimostrazione. Basta applicare il Teorema di Weierstrass all'insieme compatto K; su tale insieme esistono massimo e minimo. In particolare, se $x_0 \in K$ é punto di minimo, allora $f(x_0) \leq f(x)$ per ogni $x \in K$ e quindi in particolare per $x = x_1$. Quindi anche per $x \in E \setminus K$ vale che $f(x) \geq f(x_1) \geq f(x_0)$, cioè $f(x) \geq f(x_0)$ per ogni $x \in E$. Analogo ragionamento si applica nella ricerca del massimo. □

Osservazione 5.3 Nel caso in cui E è un insieme chiuso, solitamente si cerca l'insieme compatto nella forma $K = E \cap \bar{B}_r$.

L'esistenza di massimi e minimi è garantita solo sotto l'ipotesi di continuità sulla funzione e della compattezza del dominio; non occorre che la funzione sia differenziabile. Il calcolo differenziale fornisce un metodo per l'individuazione dei punti di massimo e minimo tramite l'individuazione dei punti stazionari, sia liberi che vincolati.

Proposizione 5.4 *Sia $f : E \to \mathbb{R}$ e $x_0 \in E$ e sia x_0 punto di differenziabilità per f. Se x_0 è un punto di minimo locale, allora $-\nabla f(x_0) \in \mathrm{Nor}(E, x_0)$ e quindi se x_0 è interno ad E, $\nabla f(x_0) = 0$. Se x_0 è un punto di massimo locale, allora $\nabla f(x_0) \in \mathrm{Nor}(E, x_0)$ e quindi se x_0 è interno ad E, $\nabla f(x_0) = 0$.*

Dimostrazione. Usando la definizione di differenziabilità in x_0 e il fatto che x_0 è punto di minimo locale, abbiamo che

$$f(x_0) \leq f(x) = f(x_0) + \nabla f(x_0) \cdot (x - x_0) + o(\|x - x_0\|).$$

Semplificando e dividendo per $\|x - x_0\|$ troviamo quindi che

$$-\nabla f(x_0) \cdot \frac{x - x_0}{\|x - x_0\|} + \frac{o(\|x - x_0\|)}{\|x - x_0\|} \leq 0.$$

In conclusione, se prendiamo $x = x_h \to x_0$ e

$$\hat{v} = \lim_{h \to +\infty} \frac{x - x_0}{\|x - x_0\|} \in \mathrm{Tan}(E, x_0),$$

abbiamo trovato che

$$-\nabla f(x_0) \cdot \hat{v} \leq 0,$$

cioè $-\nabla f(x_0) \in \mathrm{Nor}(E, x_0)$. Se poi x_0 è interno ad E, allora $\mathrm{Tan}(E, x_0) = \mathbb{R}^n$ e quindi $\mathrm{Nor}(E.x_0) = \{0\}$. Il caso di punto di massimo è analogo. □

In definitiva, un punto estremale x_0 interno ad E è un *punto stazionario libero*, o semplicemente *punto stazionario*, per f, cioè

$$\nabla f(x_0) = 0.$$

Per un punto estremale x_0 che appartiene alla frontiera ∂E di E si ha invece che x_0 è un *punto stazionario vincolato* per f, ,cioè

$$-\nabla f(x_0) \in \mathrm{Nor}(E, x_0) \qquad \text{se } x_0 \text{ punto di minimo,}$$

mentre

$$\nabla f(x_0) \in \mathrm{Nor}(E, x_0) \qquad \text{se } x_0 \text{ punto di massimo.}$$

Per la determinazione dei punti stazionari vincolati esistono sostanzialmente due metodi, a seconda che la frontiera ∂E di E sia nel x_0 una varietà implicita o parametrizzata. Il primo metodo viene chiamato *metodo dei moltiplicatori di Lagrange*, il secondo *metodo di parametrizzazione del vincolo*.

Teorema 5.5 (Moltiplicatori di Lagrange) *Supponiamo che x_0 sia un punto estremale per $f : E \to \mathbb{R}$ funzione differenziabile in x_0, $E \subset \mathbb{R}^n$ insieme chiuso e supponiamo si abbia $x_0 \in \Sigma \subset \partial E$ con Σ k–varietà implicita definita mediante la funzione g a valori in $\mathbb{R}^{n-k}$. Allora esiste $\lambda \in R^{n-k}$ tale che (x_0, λ) è un punto stazionario libero della funzione*

$$\mathcal{L}(x, \lambda) = f(x) - \lambda \cdot g(x) = f(x) - \lambda_1 g_1(x) + \ldots - \lambda_{n-k} g_{n-k}(x). \tag{5.2}$$

La funzione $\mathcal{L}$ viene detta funzione Lagrangiana, mentre i numeri $\lambda_1, \ldots, \lambda_{n-k}$ vengono chiamati moltiplicatori di Lagrange.

DIMOSTRAZIONE. La dimostrazione segue semplicemente notando che se $x_0 \in \Sigma$, allora necessariamente $g(x_0) = 0$, mentre dato che Σ è una k–varietà implicita, allora per quanto visto nel capitolo precedente

$$N_{x_0}\Sigma = \langle \nabla g_1(x_0), \dots, \nabla g_{n-k}(x_0)\rangle.$$

Quindi, dato che $\Sigma \subset \partial E \subset E$, allora $\mathrm{Nor}(E, x_0) \subset N_{x_0}\Sigma$ e quindi se $-\nabla f(x_0)$ o $\nabla f(x_0)$ appartiene a $\mathrm{Nor}(E, x_0)$, necessariamente devono esistere $\lambda_1, \dots, \lambda_{n-k}$ tali che

$$\nabla f(x_0) = \lambda_1 \nabla g_1(x_0) + \ldots + \lambda_{n-k} \nabla g_{n-k}(x_0).$$

Quindi necessariamente x_0 deve essere un punto stazionario libero della funzione Lagrangiana di $2n - k$ variabili definita in (5.2) $\square$

Osservazione 5.6 Dalla dimostrazione del precedente Teorema si evince che l'insieme Σ non è necessario che sia una k–varietà implicita, ma è sufficiente che sia il livello 0 di una funzione g a valori in $\mathbb{R}^{n-k}$ per cui

$$\mathrm{Nor}(\Sigma, x_0) \subset \langle \nabla g_1(x_0), \dots, \nabla g_{n-k}(x_0)\rangle.$$

Se ad esempio vogliamo determinare i punti di massima e minima distanza dall'origine dei punti dell'insieme

$$\Sigma = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 = 1, x^2 + z^2 = 1\}.$$

introducendo la funzione Lagrangiana con la funzione quadrato della distanza dall'origine $f(x, y, z) = x^2 + y^2 + z^2$

$$\mathcal{L}(x, y, z, \lambda_1, \lambda_2) = x^2 + y^2 + z^2 - \lambda_1(x^2 + y^2 - 1) - \lambda_2(x^2 + z^2 - 1),$$

troviamo i punti stazionari per $(\pm 1, 0, 0)$ con $(\lambda_1, \lambda_2) = (\alpha, 1 - \alpha)$, $\alpha \in \mathbb{R}$, e $(0, \pm 1, \pm 1)$ con $(\lambda_1, \lambda_2) = (1, 1)$. Per confronto nei valori si trova che Σ ha minimi distanza 1 dall'origine e i punti di minima distanza sono $(1, 0, 0)$ e $(-1, 0, 0)$, mentre la massima ha massima distanza $\sqrt{2}$ dall'origine e i quattro punti $(0, 1, 1)$, $(0, -1, 1)$, $(0, 1, -1)$ e $(0, -1, -1)$ sono i punti di massima distanza. Nei punti di minima distanza però Σ non è localmente una curva ma il metodo dei moltiplicatori di Lagrange in tali punti si applica perchè

$$\mathrm{Tan}(\Sigma, (\pm 1, 0, 0)) = \langle (0, 1, 1)\rangle \cup \langle (0, 1 - 1)\rangle,$$

e quindi

$$\mathrm{Nor}(\Sigma, (\pm 1, 0, 0)) = \langle (1, 0, 0)\rangle.$$

Per quanto riguarda la funzione

$$g(x, y, z) = (x^2 + y^2 - 1, x^2 + z^2 - 1),$$

il suo Jacobiano in $(\pm 1, 0, 0)$ è dato da

$$Jg(\pm 1, 0, 0) = \begin{pmatrix} \pm 2 & 0 & 0 \\ \pm 2 & 0 & 0 \end{pmatrix}$$

e tale matrice ha rango 1, però resta vero che i vettori $\nabla g_1(\pm 1, 0, 0) = \nabla g_2(\pm 1, 0, 0) = (\pm 2, 0, 0)$ generano il normale a Σ in $(\pm 1, 0, 0)$.

Mostriamo anche il metodo di parametrizzazione del vincolo descritto dal seguente risultato.

Teorema 5.7 (Parametrizzazione del vincolo) *Sia x_0 un punto estremale vincolato per la funzione $f : E \to \mathbb{R}$ differenziabile in x_0, E insieme chiuso; supponiamo $x_0 \in \Sigma \subset \partial E$ con Σ una k–varietà parametrizzata mediante $r : D \to \mathbb{R}^n$, $D \subset \mathbb{R}^k$. Se $x_0 = r(u_0)$ con u_0 punto interno per D, allora u_0. Allora u_0 è un punto stazionario libero della funzione $h(u) = f(r(u))$.s*

Dimostrazione. Basta notare che $\Sigma \subset \partial E \subset E$ e quindi $\mathrm{Nor}(E, x_0) \subset N_{x_0}\Sigma$; la richiesta che $-\nabla f(x_0)$ o $\nabla f(x_0)$ appartenga a $\mathrm{Nor}(E, x_0)$ implica che $\nabla f(x_0)$ è ortogonale a tutti i vettori tangenti a Σ in x_0, quindi ai vettori $r_{u_i}(u_0)$. Abbiamo quindi che

$$\nabla f(x_0) \cdot r_{u_i}(u_0) = 0, \qquad \forall i = 1, \ldots, k.$$

Se definiamo quindi $h = f \circ r$ troviamo che

$$\frac{\partial h}{\partial u_i}(u) = \nabla f(r(u)) \cdot r_{u_i}(u), \qquad \forall i = 1, \ldots, k$$

che quindi per $u = u_0$ implica che $\nabla h(u_0) = 0$, cioè u_0 è un punto stazionario libero per h. □

Chiudiamo questa sezione con un esempio riassuntivo di determinazione di massimo e minimo su un insieme compatto.

Esempio 5.1 Determinare massimo e minimo della funzione

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

sull'insieme

$$E = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 \le z^2, 0 \le z \le 2\}.$$

L'insieme dato ha una parte interna

$$E^\circ = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 < z^2, 0 < z < 2\}$$

e una frontiera che possiamo decomporre come

$$\partial E = \Sigma_1^{(2)} \cup \Sigma_2^{(2)} \cup \Sigma^{(1)} \cup \Sigma^{(0)},$$

dove

$$\Sigma_1^{(2)} = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 = z^2, 0 < z < 2\},$$

$$\Sigma_2^{(2)} = \{(x, y, z) \in \mathbb{R}^3 : z = 2, x^2 + y^2 < z^2\},$$

$$\Sigma^{(1)} = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 = z^2, z = 2\}, \qquad \Sigma^{(0)} = \{(0, 0, 0)\}.$$

Iniziamo col cercare i punti stazionari liberi della funzione f; essi sono determinati dalla condizione

$$\nabla f(x, y, z) = (3x^2, 3y^2, -2z) = (0, 0, 0);$$

si trova l'unico punto stazionario $(0, 0, 0)$ che però non è interno ad E.

L'insieme $\Sigma_1^{(2)}$ è definito come zero della funzione $g(x, y, z) = x^2 + y^2 - z^2$ con $0 < z < 2$, ma può anche essere visto come grafico della funzione $z = \sqrt{x^2 + y^2}$ con $0 < \sqrt{x^2 + y^2} < 2$. Possiamo quindi usare equivalentemente sia il metodo dei moltiplicatori di Lagrange con

$$\mathcal{L}(x, y, z, \lambda) = x^3 + y^3 - z^2 - \lambda(x^2 + y^2 - z^2)$$

che il metodo per parametrizzazione del vincolo definendo

$$h(x, y) = f(x, y, \sqrt{x^2 + y^2}) = x^3 + y^3 - x^2 - y^2.$$

Nel primo caso troviamo che i punti stazionari per $\mathcal{L}$ sono dati da $(0,0,0)$ con λ arbitrario, $(3/2,0,\pm 3/2)$, $(0,3/2,\pm 3/2)$ e $(3/2,3/2,\pm 3\sqrt{2}/2)$ con $\lambda = 1$. Tra questi quelli che soddisfano la condizione $0 < z < 2$ sono $(3/2,0,3/2)$, $(0,3/2,3/2)$ e $(3/2,3/2,3\sqrt{2}/2)$. Nel secondo caso invece troviamo i punti stazionari di h dati da $(0,0)$, $(3/2,0)$, $(0,3/2)$ e $(3/2,3/2)$; la codizione $0 < \sqrt{x^2+y^2} < 2$ esclude il primo punto stazionario.

L'insieme $\Sigma_2^{(2)}$ è definito come lo zero della funzione $g(x,y,z) = z-2$ con $\sqrt{x^2+y^2} < 2$, ma può anche essere visto come il grafico della funzione $z = 2$ con $\sqrt{x^2+y^2} < 2$. Se usiamo il metodo dei moltiplicatori di Lagrange definiamo

$$\mathcal{L}(x,y,z,\lambda) = x^3 + y^3 - z^2 - \lambda(z-2),$$

mentre usando il metodo di parametrizzazione del vincolo si definisce

$$h(x,y) = f(x,y,2) = x^3 + y^3 - 4.$$

Nel primo caso si trova un solo puto stazionario per $\mathcal{L}$ dato da $(0,0,4)$ con $\lambda = 4$ che soddisfa la condizione $\sqrt{x^2+y^2} < 2$. Nel secondo caso h ha un solo punto stazionario in $(0,0)$.

L'insieme $\Sigma^{(1)}$ è definito come zero della funzione vettoriale

$$g(x,y,z) = (x^2 + y^2 - z^2, z-2);$$

analogamente, ponendo $z = 2$ in $x^2+y^2 = z^2$, cioè scrivendo $z = 2$ e $x^2+y^2 = 4$ si riconosce una circonferenza di raggio 2 nel piano $z = 2$, che può essere parametrizzzata da

$$r(t) = (2\cos t, 2\sin t, 2), \qquad t \in [0, 2\pi].$$

Possiamo quindi considerare il metodo dei moltiplicatori di Lagrange definendo

$$\mathcal{L}(x,y,z,\lambda_1,\lambda_2) = x^3 + y^3 - z^2 - \lambda_1(x^2+y^2-z^2) - \lambda_2(z-2),$$

o equivalentemente si può usare il metodo di parametrizzazione del vincolo ponendo

$$h(t) = f(r(t)) = 8\cos^3 t + 8\sin^3 t - 4.$$

Nel primo metodo troviamo i punti stazionari di $\mathcal{L}$ che sono dati da $(2,0,2)$ e $(0,2,2)$ con $(\lambda_1,\lambda_2) = (3,8)$, $(-2,0,2)$ e $(0,-2,2)$ con $(\lambda_1,\lambda_2) = (-3,-16)$, $(\sqrt{2},\sqrt{2},2)$ con $(\lambda_1,\lambda_2) = (3\sqrt{2}/2, 6\sqrt{2}-4)$ e $(-\sqrt{2},-\sqrt{2},2)$ con $(\lambda_1,\lambda_2) = (-3\sqrt{2}/2, -6\sqrt{2}-4)$; tutti questi punti sono punti stazionari di f vincolati ad E. Nel secondo metodo troviamo che i punti stazionari di h sono dati da 0, $\pi/4$, $\pi/2$, π, $5\pi/4$ e $3\pi/2$.

Resta l'insieme $\Sigma^{(0)} = \{(0,0,0)\}$; in tale insieme non ha senso considerare il calcolo differenziale e quindi l'unica cosa da fare è valutare f in tale punto.

In definitiva, valutando f in tutti i punti stazionari, sia liberi che vincolati, che abbiamo trovato, per confronto sui valori si trova che

$$\min_E f = -12, \qquad \text{assunto in } (-2,0,2) \text{ e } (0,-2,2),$$

mentre

$$\max_E f = 4, \qquad \text{assunto in } (2,0,2) \text{ e } (0,2,2).$$

Osservazione 5.8 Eventuali punti di non differenziabilità per una funzione continua diventano automaticamente candidati punti di massimo e minimo. Prendiamo ad esempio la funzione

$$f(x,y) = \sqrt{x^2+y^2} + x^2$$

e cerchiamone massimo e minimo sul $\bar{B}_1(0,0)$. La funzione non è differenziabile in $(0,0)$ dove vale 0 e non presenta punti stazionari liberi. I punti stazionari vincolati su $x^2+y^2=1$ possono essere cercati parametrizzando il vincolo con $r(t) = (\cos t, \sin t)$; si trovano 4 punti stazionari in $(1,0)$, $(0,1)$, $(-1,0)$ e $(0,-1)$. Nel primo e terzo punto la funzione vale 2, negli altri due vale 1. Pertanto il massimo della funzione è 2 e $(1,0)$ e $(-1,0)$ sono punti di massimo, mentre il minimo è 0 e $(0,0)$ è punto di minimo.

5.2 Punti stazionari liberi e loro classificazione

Trattiamo in questa sezione un problema diverso rispetto alla determinazione dei massimi e minimi assoluti delle funzioni differenziabili; la classificazione dei punti stazionari liberi.

Abbiamo visto la definizione di punto stazionario libero e abbiamo anche visto la relazione tra punti stazionari e punti di massimo e minimo locali. Classificare un punto stazionario x_0 significa determinare quale delle seguenti tre alternative è verificata:

1. x_0 punto di minimo locale o minimo locale stretto;
2. x_0 punto di massimo locale o massimo locale stretto;
3. x_0 punto di sella, cioè un punto per il quale per ogni $r > 0$ la funzione f ammette su $E \cap B_r(x_0)$ valori sia maggiori che minori di $f(x_0)$ e quindi tale che eistono $x_1, x_2 \in E \cap B_r(x_0)$ tali che
$$f(x_1) < f(x_0) < f(x_2).$$

5.2.1 Forme quadratiche

Faremo la classificazione dei punti stazionari per funzioni di classe C^2 e lo strumento principale sarà la Formula di Taylor del secondo ordine. Prima di procedere ricordiamo le seguenti definizioni. Ricordiamo che una forma quadratica in $\mathbb{R}^n$ è definita da una funzione $Q : \mathbb{R}^n \to \mathbb{R}$ tale che

$$Q(\lambda v) = \lambda^2 v, \qquad \forall v \in \mathbb{R}^n, \forall \lambda \in \mathbb{R}$$

e

$$B(u,v) = \frac{1}{2}\left(Q(u+v) - Qu - Qv\right) \tag{5.3}$$

è bilineare.

Definizione 5.9 *Sia $Q : \mathbb{R}^n \to \mathbb{R}$ una forma quadratica. Diremo che:*

1. *Q è definita positiva se*
$$Qv > 0, \qquad \forall v \in \mathbb{R}^n \setminus \{0\};$$
2. *Q è semi–definita positiva se*
$$Qv \geq 0, \qquad \forall v \in \mathbb{R}^n \setminus \{0\}$$
ed esiste $v \in \mathbb{R}^n \setminus \{0\}$ per cui $Qv = 0$;

3. *Q è definita negativa se*
$$Qv < 0, \qquad \forall v \in \mathbb{R}^n \setminus \{0\};$$

4. *Q è semi–definita negativa se*
$$Qv \leq 0, \qquad \forall v \in \mathbb{R}^n \setminus \{0\}$$
ed esiste $v \in \mathbb{R}^n \setminus \{0\}$ per cui $Qv = 0$;

5. *Q è indefinita se esistono $v_1, v_2 \in \mathbb{R}^n \setminus \{0\}$ tali che*
$$Qv_1 < 0 < Qv_2.$$

Lo studio della definizione di una forma quadratica può essere ridotta allo studio della sua definizione per $v \in \mathbb{S}^{n-1}$ grazie al fatto che
$$Qv = \|v\|^2 Q\frac{v}{\|v\|}, \qquad \forall v \in \mathbb{R}^n \setminus \{0\}.$$
Inoltre, se A è la matrice simmetrica associata alla forma bilineare B definita da Q, allora possiamo classificare Q mediante la matrice A nel seguente modo.

Proposizione 5.10 *Sia $A \in \mathbb{R}^{n\times n}$ la matrice simmetrica associata alla forma bilineare B definita dalla forma quadratica Q tramite* (5.3). *Se denotiamo con $\lambda)_1, \ldots, \lambda_n$ gli autovalori A. Allora:*

1. *Q è definita positiva se e solo se $\lambda_i > 0$ per ogni $i = 1, \ldots, n$;*
2. *Q è semi–definita positiva se e solo se $\lambda_i \geq 0$ per ogni $i = 1, \ldots, n$ ed esiste almeno un autovalore $\lambda_{i_0} = 0$;*
3. *Q è definita negativa se e solo se $\lambda_i < 0$ per ogni $i = 1, \ldots, n$;*
4. *Q è semi–definita negativa se e solo se $\lambda_i \leq 0$ per ogni $i = 1, \ldots, n$ ed esiste almeno un autovalore $\lambda_{i_0} = 0$;*
5. *Q è indefinita se esiste un autovalore $\lambda_{j_0} < 0$ ed un autovalore $\lambda_{i_0} < 0$.*

Dimostrazione. La dimostrazione si basa sul fatto che una matrice simmetrica A ammette una base ortonormale $v_1, \ldots, v_n \in \mathbb{S}^{n-1}$ di autovettori con rispettivi autovalori $\lambda_1, \ldots, \lambda_n$; quindi possiamo scrivere un generico elemento $v \in \mathbb{R}^n$ come combinazione
$$v = \sum_{i=1}^{N} c_i v_i = c_1 v_1 + \ldots + c_n v_n,$$
da cui
$$Qv = A \cdot v \cdot v = \sum_{i=1}^{n} c_i^2 \lambda_i. \tag{5.4}$$
Prendendo nella precedente espressione $v = v_i$ si trova quindi che $Qv_i = \lambda_i$ e quindi se Q è semidefiinita positiva allora $\lambda_i \geq 0$ per ogni $i = 1, \ldots, n$, definita positiva allora $\lambda_i > 0$. Analogo discorso vale per il caso semidefinito negativo e definito negativo.

Se poi eistono due vettori $z, w \in \mathbb{R}^n$ per cui $Qz < 0 < Qw$, allora dall'espressione (5.4) denotando con $\lambda_{\min}$ e $\lambda_{\max}$ rispettivamente il più piccolo e il più grande autovalore, otteniamo che
$$\lambda_{\min}\|z\|^2 \leq \lambda_1 z_1^2 + \ldots + \lambda_n z_n^2 = Qz < 0 < Qw = \lambda_1 w_1^2 + \ldots + \lambda_n w_n^2 \leq \lambda_{\max}\|w\|^2$$

ricaviamo che $\lambda_{\min} < 0 < \lambda_{\max}$.

Viceversa, se tutti gli autovalori sono non negativi, da (5.4) ricaviamo che $Qv \geq 0$ essendo somma di termini tutti non negativi. Se esiste un autovalore nullo $\lambda_{i_0} = 0$, allora $Qv_{i_0} = 0$ e quindi Q è semidefinita positiva. Se gli autovalori sono tutti strettamente posivi, allora dato che $v \neq 0$, almeno uno degli addendi a destra di (5.3) è non nullo da cui $Qv > 0$, cioè Q definita positivas. Analogo discorso si fa nel caso di autovalori non positivi o strettamente negativi. Infine se esistono $\lambda_{j_0} < 0 < \lambda_{i_0}$, allora

$$Qv_{j_0} = \lambda_{j_0} < 0 < \lambda_{i_0} = Qv_{i_0},$$

cioè Q è indefinita. □

La determinazione degli autovalori è solitamente un problema facile nel caso $n = 2$ in quanto basta trovare le radici di un polinomio di secondo grado. Diventa molto compilcato invece nel caso $n \geq 3$. Osserviamo anche che per studiare la definizione di una forma quadratica non serve sapere esattamente il valore degli autovalori, ma basta sapere il loro segno. Uno studio completo del segno degli autovalori nel caso $n = 2$ può essere fatto mediante determinante e traccia.

Proposizione 5.11 *Sia $A \in \mathbb{R}^{2\times 2}$ una matrice simmetrica. Allora:*

1. *A è semi–definita positiva se e solo se* $\det A = 0$ *e* $\mathrm{tr} A \geq 0$*;*
2. *A è definita positiva se e solo se* $\det A > 0$ *e* $\mathrm{tr} A > 0$*;*
3. *A è semi–definita negativa se e solo se* $\det A = 0$ *e* $\mathrm{tr} A \leq 0$*;*
4. *A è definita negativa se e solo se* $\det A > 0$ *e* $\mathrm{tr} A < 0$*;*
5. *A è indefinita se* $\det A < 0$.

Dimostrazione. Basta notare che nel caso bidimensionale $\det A = \lambda_1\lambda_2$ e $\mathrm{tr} A = \lambda_1 + \lambda_2$. Se $\det A = 0$ allora un autovalore è nullo e la traccia coincide con l'altro autovalore. Quindi la matrice è semidefinita, positiva se il secondo autovalore è non negativo, negativa se non positivo. Se il determinante è positivo allora i due autovalori sono entrambi non nulli e di segno concorde; sono positivi se e solo sse la traccia è positiva, negativi altrimenti. Infine se il determinante è negativo, allora i due autovalori sono non nulli e di segno discorde. □

In dimensione $n \geq 3$ determinante e traccia non sono più sufficienti per uno studio completo ma possono dare solo informazioni parziali. Per questo motivo presentiamo il seguente risultato. Data una matrice $A \in \mathbb{R}^{n\times n}$, si definisce *minore principale di testa* o *di nord ovest* di A di ordine k, $k = 1, \dots, n$, la matrice quadrata $A^{(k)} \in \mathbb{R}^{k\times k}$, ottenuta selezionando le prime k righe e le prime k colonne della matrice A. Quindi ad se

$$A = \begin{pmatrix} 3 & 0 & 1 \\ 0 & 2 & 4 \\ 1 & 4 & 5 \end{pmatrix},$$

allora

$$A^{(1)} = 3, \qquad A^{(2)} = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}, \qquad A^{(3)} = A.$$

Teorema 5.12 (Sylvester) *Sia $A \in \mathbb{R}^{n\times n}$ una matrice simmetrica. Allora:*

1. *A è definita positiva se e solo se*

$$\det A^{(k)} > 0, \qquad \forall k = 1, \dots, n; \tag{5.5}$$

2. *A è definita negativa se e solo se*

$$(-1)^k \det A^{(k)} > 0, \qquad \forall k = 1, \ldots, n. \tag{5.6}$$

Il Teorema di Sylvester caratterizza le matrici definite, o positive o negative. Si possono caratterizzare le matrici semi–definite come enunciato nel teorema che segue. Denotiamo in tale teorema con $A_p^{(k)}$ il minore principale ottenuto da A selezionando k righe e k colonne aventi lo stesso indice. Quindi ad esempio se

$$A = \begin{pmatrix} 3 & 0 & 1 \\ 0 & 2 & 4 \\ 1 & 4 & 5 \end{pmatrix},$$

allora abbiamo tre possibili $A_p^{(1)}$ che sono gli elementi 3, 2 e 5 della diagonale di A, abbiamo tre possibili scelte per $A_p^{(2)}$ che sono date da

$$\begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}, \begin{pmatrix} 3 & 1 \\ 1 & 5 \end{pmatrix}, \begin{pmatrix} 2 & 4 \\ 4 & 5 \end{pmatrix},$$

mentre $A_p^{(3)} = A$.

Teorema 5.13 *La matrice simmetrica $A \in \mathbb{R}^{n \times n}$ è semi–definita positiva se e solo se*

$$\det A_p^{(k)} \geq 0$$

per ogni minore principale di ordine k e per ogni $k = 1, \ldots, n$. Avremo invece che A è semi–definita negativa se e solo se

$$(-1)^k \det A_p^{(k)} \geq 0$$

per ogni minore principale di ordine k e per ogni $k = 1, \ldots, n$.

Se non sono soddisfatte le condizioni dei due teoremi appena enunciati avremo quindi che la matrice A è indefinita.

Riassumendo, se $A \in \mathbb{R}^{2 \times 2}$ è una matrice simmetrica

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

con $a_{12} = a_{21}$, avremo che:

1. A è semi–definita positiva se e solo se

$$a_{11} \geq 0, \quad a_{22} \geq 0, \quad a_{11}a_{22} - a_{12}^2 \geq 0 :$$

2. A è definita positiva se e solo se

$$a_{11} > 2, \quad a_{11}a_{22} - a_{12}^2 > 0;$$

3. A è semi–definita negativa se e solo

$$a_{11} \leq 0, \quad a_{22} \leq 0, \quad a_{11}a_{22} - a_{12}^2 \geq 0 :$$

4. A è definita negativa se e solo se

$$a_{11} < 2, \quad a_{11}a_{22} - a_{12}^2 > 0;$$

5. A è indefinita se e solo se

$$a_{11}a_{22} - a_{12}^2 < 0.$$

Per una matrice simmetrica $A \in \mathbb{R}^{3\times 3}$

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

con $a_{12} = a_{21}$, $a_{13} = a_{31}$ e $a_{23} = a_{32}$, avremo invece che

1. A è semi–definita positiva se e solo se

$$a_{11} \geq 0, \quad a_{22} \geq 0, \quad a_{33} \geq 0,$$
$$a_{11}a_{22} - a_{12}^2 \geq 0, \quad a_{11}a_{33} - a_{13}^2 \geq 0, \quad a_{22}a_{33} - a_{23}^2 \geq 0,$$
$$a_{11}a_{22}a_{33} + 2a_{12}a_{23}a_{13} - a_{22}a_{13}^2 - a_{11}a_{23}^2 - a_{33}a_{12}^2 \geq 0;$$

2. A è definita positiva se e solo se

$$a_{11} > 2, \quad a_{11}a_{22} - a_{12}^2 > 0, \quad a_{11}a_{22}a_{33} + 2a_{12}a_{23}a_{13} - a_{22}a_{13}^2 - a_{11}a_{23}^2 - a_{33}a_{12}^2 > 0;$$

3. A è semi–definita negativa se e solo se

$$a_{11} \leq 0, \quad a_{22} \leq 0, \quad a_{33} \leq 0,$$
$$a_{11}a_{22} - a_{12}^2 \geq 0, \quad a_{11}a_{33} - a_{13}^2 \geq 0, \quad a_{22}a_{33} - a_{23}^2 \geq 0,$$
$$a_{11}a_{22}a_{33} + 2a_{12}a_{23}a_{13} - a_{22}a_{13}^2 - a_{11}a_{23}^2 - a_{33}a_{12}^2 \leq 0;$$

4. A è definita negativa se e solo se

$$a_{11} < 2, \quad a_{11}a_{22} - a_{12}^2 > 0, \quad a_{11}a_{22}a_{33} + 2a_{12}a_{23}a_{13} - a_{22}a_{13}^2 - a_{11}a_{23}^2 - a_{33}a_{12}^2 < 0;$$

5. A è indefinita negli altri casi.

5.2.2 Classificazione dei punti stazionari

Mettiamo ora in relazione le proprietà delle forme quadratiche con la classificazione dei punti stazionari.

Lo strumento fondamentale nel seguito sarà la formula di Taylor del secondo ordine centrata in x_0 per una funzione loclamente di classe C^2 attorno al punto x_0

$$f(x) = f(x_0) + \nabla f(x_0) \cdot (x - x_0) + \frac{1}{2} Hf(x_0)(x - x_0) \cdot (x - x_0) + o(\|x - x_0\|^2). \tag{5.7}$$

Abbiamo il seguente risultato:

Proposizione 5.14 *Sia $f : E \to \mathbb{R}$ una funzione localmente di classe C^2 in un intorno del punto x_0 interno ad E. Si hanno le seguenti condizioni necessarie:*

1. *se x_0 è un punto di minimo locale, allora $\nabla f(x_0) = 0$ e $Hf(x_0)$ è semi–definita positiva;*
2. *se x_0 è un punto di massimo locale, allora $\nabla f(x_0) = 0$ e $Hf(x_0)$ è semi–definita negativa.*

Per un punto stazionario x_0 avremo invece le seguenti condizioni suffcienti:

1. *se $Hf(x_0)$ è definita positiva, allora x_0 è un punto di minimo locale stretto;*
2. *se $Hf(x_0)$ è definita negativa, allora x_0 è un punto di massimo locale stretto;*
3. *se $Hf(x_0)$ è indefinita, allora x_0 è un punto di sella.*

Dimostrazione. Per quanto riguarda le condizioni sufficienti, abbiamo già visto che un punto estremale è un punto stazionario e quindi $\nabla f(x_0) = 0$. Se ad esempio x_0 è un punto di minimo locale, allora dalla formula di Taylor (5.7) deduciamo che esiste $r > 0$ tale che per ogni $x \in B_r(x_0) \subset E$

$$0 \leq f(x) - f(x_0) = \frac{1}{2} Hf(x_0)(x - x_0) \cdot (x - x_0) + o(\|x - x_0\|^2).$$

Scrivendo $x = x_0 + tv$ con $|t| < r$, $v \in \mathbb{S}^{n-1}$, deduciamo che

$$0 \leq f(x_0 + tv) - f(x_0) = \frac{t^2}{2} Hf(x_0)v \cdot v + o(t^2);$$

dividendo per $t^2/2$ e passando al limite $t \to 0$ troviamo quindi che

$$Hf(x_0) \cdot v \cdot v \geq 0$$

e quindi $Hf(x_0)$ é semi–definita positiva. In modo analogo si ragiona nel caso di massimo locale.

Veniamo ora alle condizioni sufficienti; denotiamo con $\lambda_{\min}$ e $\lambda_{\max}$ il più piccolo e il più grande autovalore di $Hf(x_0)$. La matrice sarà definita positiva se e solo se $\lambda_{\min} > 0$, definita negativa se e solo se $\lambda_{\max} < 0$ e indefinita se e solo se $\lambda_{\min} < 0 < \lambda_{\max}$.
Supponiamo quindi che x_0 sia punto stazionario e che $Hf(x_0)$ sia definita positiva: dalla formula di Taylor (5.7) troviamo che

$$\begin{aligned} f(x) - f(x_0) =& \frac{1}{2} Hf(x_0) \cdot (x - x_0) \cdot (x - x_0) + o(\|x - x_0\|^2) \geq \frac{\lambda_{\min}}{2} \|x - x_0\|^2 + o(\|x - x_0\|^2) \\ =& \|x - x_0\|^2 \left(\frac{\lambda_{\min}}{2} + \frac{o(\|x - x_0\|^2)}{\|x - x_0\|^2} \right). \end{aligned}$$

Dato che

$$\lim_{x \to x_0} \frac{o(\|x - x_0\|^2)}{\|x - x_0\|^2} = 0,$$

esiste $r > 0$ per cui se $x \in B_r(x_0)$ allora

$$-\frac{\lambda_{\min}}{4} < \frac{o(\|x - x_0\|^2)}{\|x - x_0\|^2} < \frac{\lambda_{\min}}{4}.$$

In definitiva per $x \in B_r(x_0) \setminus \{x_0\}$ avremo che

$$f(x) - f(x_0) > \frac{\lambda_{\min}}{4} \|x - x_0\|^2 > 0.$$

Quindi x_0 è un punto di minimo locale stretto. Ragionamento analogo dimostra che se $Hf(x_0)$ è definita negativa, allora x_0 è punto di massimo locale stretto.
Supponiamo ora che $Hf(x_0)$ sia indefinita; sia $v_1 \in \mathbb{S}^{n-1}$ autovettore associato all'autovalore $\lambda_{\min}$ e $v_2 \in \mathbb{S}^{n-1}$ autovettore associato all'autovalore $\lambda_{\max}$. Fissiamo $r > 0$ tale per cui per $x \in B_r(x_0)$

$$-\frac{\lambda_{\max}}{4} < \frac{o(\|x - x_0\|^2)}{\|x - x_0\|^2} < -\frac{\lambda_{\min}}{4}.$$

Possiamo quindi concludere che se $0 < |t| < r$

$$f(x_0 + tv_1) - f(x_0) = \frac{1}{2} Hf(x_0) \cdot tv_1 \cdot tv_1 + o(t^2) = \frac{\lambda_{\min}}{2} t^2 + o(t^2) = t^2 \left(\frac{\lambda_{\min}}{2} + \frac{o(t^2)}{t^2} \right) < t^2 \frac{\lambda_{\min}}{4} < 0$$

mentre

$$f(x_0 + tv_2) - f(x_0) > t^2 \frac{\lambda_{\max}}{4} > 0,$$

da cui il fatto che per $0 < |t| < r$

$$f(x_0 + tv_1) < f(x_0) < f(x_0 + tv_2).$$

Quindi x_0 è un punto di sella. □

5.3 Funzioni convesse

Una importante classe di funzioni per cui i problemi di massimo e minimo possono semplificarsi è dato dalle funzioni convesse.

Definizione 5.15 (Funzione convessa) *Una funzione $f : E \to \mathbb{R}$ con $E \subset \mathbb{R}^n$ insieme convesso si dice convessa se per ogni $x_1, x_2 \in E$ e per ogni $t \in [0, 1]$*

$$f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2).$$

La funzione si dice strettamente convessa se per ogni $x_1, x_2 \in E$ e per ogni $t \in (0, 1)$

$$f(tx_1 + (1-t)x_2) < tf(x_1) + (1-t)f(x_2).$$

Se definiamo l'epigrafico di f ponendo

$$\mathrm{Epi}(f, E) = \{(x, t) \in E \times \mathbb{R} : t \geq f(x)\} \subset \mathbb{R}^{n+1},$$

la convessità di una funzione può essere caratterizzata mediante il seguente risultato.

Proposizione 5.16 *Sia $f : E \to \mathbb{R}$ con $E \subset \mathbb{R}^n$ insieme convesso. Allora f è convessa se e solo se $\mathrm{Epi}(f, E)$ è un insieme convesso.*

Dimostrazione. Supponiamo f sia convessa e prendiamo $y_1, y_2 \in \mathrm{Epi}(f, E)$. Scrivendo $y_i = (x_i, s_i)$, $i = 1, 2$, per definizione di epigrafico avremo che $s_i \geq f(x_i)$. Dobbiamo mostrare che $[y_1, y_2] \subset \mathrm{Epi}(f, E)$. Avremo che y appartiene al segmento $[y_1, y_2]$ se esiste $t \in [0, 1]$ e quindi scrivendo $y = (x, s)$,

$$y = (x, s) = ty_1 + (1-t)y_2 = (tx_1 + (1-t)x_2, ts_1 + (1-t)s_2).$$

Ma allora per la convessità dii f

$$s = ts_1 + (1-t)s_2 \geq tf(x_1) + (1-t)f(x_2) \geq f(tx_1 + (1-t)x_2) = f(x),$$

e quindi $y = (x, s) \in \mathrm{Epi}(f, E)$.

Viceversa, se prendiamo $y_1 = (x_1, f(x_1))$ e $y_2 = (x_2, f(x_2))$, sapendo che $[y_1, y_2] \subset \mathrm{Epi}(f, E)$ abbiamo che per ogni $t \in [0, 1]$ vale che $ty_1 + (1-t)y_2 = (tx_1 + (1-t)x_2, tf(x_1) + (1-t)f(x_2)) \in \mathrm{Epi}(f, E)$, quindi

$$tf(x_1) + (1-t)f(x_2) \geq f(tx_1 + (1-t)x_2),$$

da cui la convessità di f. □

Possiamo caratterizzare la convessità di una funzione con il calcolo differenziale nel seguente modo.

Teorema 5.17 *Sia $f : E \to \mathbb{R}$ con $E \subset \mathbb{R}^n$ insieme convesso. Allora:*

1. *se $f \in C^1(E)$, allora f è convessa se e solo se per ogni $x_0 \in E$*

$$f(x) \geq f(x_0) + \nabla f(x_0) \cdot (x - x_0), \qquad \forall x \in E; \tag{5.8}$$

2. *se $f \in C^1(E)$, allora f è strettamente convessa se e solo se per ogni $x_0 \in E$*

$$f(x) > f(x_0) + \nabla f(x_0) \cdot (x - x_0), \qquad \forall x \in E \setminus \{x_0\};$$

3. *se $f \in C^2(E)$, allora f è convessa se e solo se per ogni $x_0 \in E$ la matrice Hessiana $Hf(x_0)$ è semi–definita positiva;*

4. *se $f \in C^2(E)$, allora se per ogni $x_0 \in E$ la matrice Hessiana è definita positiva, f è strettamente convessa.*

Dimostrazione. Fissiamo $x_0, x \in E$ e usiamo la definizione di convessità con $x_1 = x$ e $x_2 = x_0$; troviamo quindi che per ogni $t \in (0, 1]$

$$f(x_0 + t(x - x_0)) = f(tx + (1 - t)x_0) \leq tf(x) + (1 - t)f(x_0) = f(x_0) + t(f(x) - f(x_0)).$$

Possiamo riscrivere la precedente condizione come

$$\frac{f(x_0 + t(x - x_0)) - f(x_0)}{t} \leq f(x) - f(x_0).$$

Passando al limite $t \to 0$ troviamo quindi che

$$\nabla f(x_0) \cdot (x - x_0) \leq f(x) - f(x_0),$$

cioè

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

Se f è strettamente convessa, non possiamo sostituire direttamente le disuguaglianze con disuguaglianze strette perchè queste non passano al limite; supponiamo quindi che si abbia $x \neq x_0$ per cui

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

Allora, usando (5.8), per $t \in (0, 1)$ dovremmo avere

$$\begin{aligned} f(tx + (1 - t)x_0) <&tf(x) + (1 - t)f(x_0) = f(x_0) + t\nabla f(x_0) \cdot (x - x_0) = f(x_0) + \nabla f(x_0) \cdot (t(x - x_0)) \\ =&f(x_0) + \nabla f(x_0) \cdot (tx + (1 - t)x_0) - x_0) \leq f(tx + (1 - t)x_0) \end{aligned}$$

e quindi un assurdo.
Viceversa, fissiamo $x_1, x_2 \in E$. Allora per ogni $x_0 \in E$

$$f(x_1) \geq f(x_0) + \nabla f(x_0) \cdot (x_1 - x_0) \tag{5.9}$$

$$f(x_2) \geq f(x_0) + \nabla f(x_0) \cdot (x_2 - x_0). \tag{5.10}$$

Se $x_0 = tx_1 + (1 - t)x_2$, allora $x_1 - x_0 = (1 - t)(x_1 - x_2)$ e $x_2 - x_0 = t(x_2 - x_1)$; moltiplicando quindi (5.9) per t e (5.10) per $(1 - t)$ e sommando le due disuguaglianze troviamo quindi che

$$tf(x_1) + (1 - t)f(x_2) \geq f(tx_1 + (1 - t)x_2),$$

cioè la convessità di f. Se valgono le disuguaglianze strette si ricava subito la stretta convessità.

Nel caso di $f \in C^2(E)$, fissiamo $x_0 \in E$ allora sappiamo che

$$f(x) \geq f(x_0) + \nabla f(x_0) \cdot (x - x_0), \qquad \forall x \in E.$$

Usando la formula di Taylor (5.7) otteniamo quindi che

$$\frac{1}{2}Hf(x_0) \cdot (x - x_0) \cdot (x - x_0) + o(\|x - x_0\|^2) \geq 0.$$

Prendendo $x = x_0 + tv$ con $v \in \mathbb{S}^{n-1}$, dalla relazione precedente otteniamo che

$$\frac{t^2}{2} Hf(x_0) \cdot v \cdot v + o(t^2) \geq 0.$$

Dividendo per $t^2/2$ e passando al limite $t \to 0$ arriviamo quindi a

$$Hf(x_0) \cdot v \cdot v \geq 0, \qquad \forall v \in \mathbb{S}^{n-1}$$

e quindi $Hf(x_0)$ è semi–definita positiva. Viceversa, se l'Hessiana è semidefinita positiva per ogni $x \in E$, fissati $x_0, x \in E$, usando la formula di Taylor con resto di Lagrange alla funzione $g(t) = f(x_0 + tx)$ troviamo che

$$f(x) = f(x_0) + \nabla f(x_0) \cdot (x - x_0) + \frac{1}{2} Hf(\xi) \cdot (x - x_0) \cdot (x - x_0)$$

con $\xi \in [x_0, x]$. Quindi se la matrice Hessiana è semi–definita positiva in ogni punto di E,

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

Grazie al punto precedente, questo implica la convessità di f. Se la matrice Hessiana di f è definita positiva, la disuguaglianza precedente diventa stretta e quindi f è strettamente convessa. □

Osservazione 5.18 Si noti che nel caso C^2 non si è caratterizzata la stretta convessità con la postiva definizione della matrice Hessiana, ma abbiamo solo una condizione sufficiente per la stretta convessità. Considerando la funzione strettamente convessa $f(x) = x^4$, $x \in \mathbb{R}$, si nota infatti che ci sono funzioni strettamente convesse la cui matrice Hessiana non è definita positiva ma solo semi–definita positiva.

Corollario 5.19 *Sia $f : E \to \mathbb{R}$, $E \subset \mathbb{R}^n$ insieme convesso, una funzione covessa; se f è differenziabile in $x_0 \in E$ e x_0 punto stazionario, allora x_0 è punto di minimo assoluto per f su E. Se poi f è strettamente convesso, allora x_0 è l'unico punto di minimo.*

Dimostrazione. Il risultato segue semplicemente da

$$f(x) \geq f(x_0) + \nabla f(x_0) \cdot (x - x_0), \qquad \forall x \in E.$$

Se x_0 è punto stazionario, allora la precedente disuguaglianza implica che x_0 è punto di minimo assoluto. Se f è strettamente convessa, allora la precedente disuguaglianza è stretta per $x \in E \setminus \{x_0\}$, e quindi x_0 è l'unico punto di minimo assoluto. □

Osservazione 5.20 Il risultato precedente ci dice quindi che per le funzioni convesse, i punti stazionari sono automaticamente punti di minimo assoluto; una funzione convessa non ammette quindi punti di sellla e i punti di massimo locale e assoluto vanno cercati sulla frontiera ∂E di E.

Esempio 5.2 Consideriamo il caso particolare di funzione radiale, $f(x) = g(\|x\|)$ con $g \in C([0, +\infty)) \cap C^2((0, +\infty))$. Abbiamo visto che la matrice Hessiana di f è data da

$$Hf(x) = g''(\|x\|) \frac{x}{\|x\|} \otimes \frac{x}{\|x\|} + \frac{g'(\|x\|)}{\|x\|} \left(I_n - \frac{x}{\|x\|} \otimes \frac{x}{\|x\|} \right);$$

quindi se $v \in \mathbb{S}^{n-1}$ è tale che $v \cdot \frac{x}{\|x\|} = 0$, ricaviamo che

$$Hf(x)v = \frac{g'(\|x\|)}{\|x\|} v$$

mentre se $v = \frac{x}{\|x\|}$

$$Hf(x) \frac{x}{\|x\|} = g''(\|x\|) \frac{x}{\|x\|}.$$

Quindi gli autovalori di $Hf(x)$ sono $g''(\|x\|)$ con autovettore $\frac{x}{\|x\|}$ e $g'(\|x\|)/\|x\|$ con autovettore ogni vettore ortogonale a $\frac{x}{\|x\|}$. La funzione è quindi convessa se e solo se g' e g'' sono non negative ed è strettamente convessa se ad esempio g' e g'' sono positive.

Bibliografia

[1] Marco Bramanti and Sandro Pagani, Carlo Domenico e Salsa. *Matematica: calcolo infinitesimale e algebra lineare*. Zanichelli, 2004.

[2] Marco Bramanti and Sandro Pagani, Carlo Domenico e Salsa. *Analisi Matematica 2*. Zanichelli, 2009.

[3] Enrico Giusti. *Analisi Matematica vol 2*. Progr. Matem. Fisica Elettronica. Boringhieri, 2003.