

Analisi Matematica II, Anno Accademico 2017-2018.

Ingegneria Edile e Architettura

Vincenzo M. Tortorelli

FOGLIO DI TEORIA n. 18

PROBLEMI DI OTTIMIZZAZIONE E ANALISI DEI PUNTI STAZIONARI

SCHEMA

I°.1 : NOZIONI DI BASE

Ottimizzazione: valori estremali, successioni minimizzanti, valori di massimo e valori di minimo, punti estremali.

Studio locale: valori estremali relativi, punti di massimo e punti di minimo relativo.

I°.2 : OTTIMIZZAZIONE : CRITERI DI ESISTENZA

Il teorema di Weierstrass.

II° : OTTIMIZZAZIONE : METODI DIRETTI

Strategie ed esempi.

III° : OTTIMIZZAZIONE : METODI INDIRECTI

Protocollo generale:

1) esistenza

2.1) elenco punti degli elementi del dominio che soddisfano le condizioni necessarie,

2.2) elenco degli elementi per cui non ha senso verificare tali condizioni necessarie

3) valutazione della funzione negli elementi dei precedenti "elenchi":

il massimo [minimo] tra questi valori sarà il massimo [minimo] su tutto il dominio

A) condizioni differenziali necessarie per punti interni.

A.1 condizioni necessarie del primo ordine. Punti stazionari.

A.2 condizioni necessarie del secondo ordine. Segnatura della matrice Hessiana in un punto critico.

B) condizioni differenziali necessarie per punti interni di una varietà con bordo.

B.1.1 condizioni necessarie del primo ordine. Punti stazionari tangenziali.

B.1.2 condizioni necessarie del primo ordine. Moltiplicatori di Lagrange.

B.2 condizioni del secondo ordine. Segnatura, della matrice Hessiana della Lagrangiana del moltiplicatore critico, sullo spazio tangente in un punto stazionario tangenziale.

C) condizioni necessarie del primo ordine per punti di frontiera regolari o di bordo.

C.1.1 Massimi e minimi relativi al bordo.

C.1.1 Segno delle derivate direzionali esterne.

- Protocollo per l'ottimizzazione usando le condizioni differenziali necessarie.

IV° : STUDIO LOCALE : CONDIZIONI DIFFERENZIALI SUFFICIENTI

A') condizioni sufficienti per punti interni.

A'.1 condizioni sufficienti di estremalità locale. Punti stazionari con Hessiana strettamente definita.

A'.2 condizioni sufficienti di estremalità locale. Punti stazionari con matrice Hessiana definita in un intorno.

A'.3 condizioni sufficienti per punti di sella.

B') condizioni sufficienti per punti stazionari tangenziali .

Segnatura stretta, della matrice Hessiana della Lagrangiana del moltiplicatore critico, sullo spazio tangente in un punto stazionario tangenziale.

I°.1 : NOZIONI DI BASE

Ottimizzazione

Valori estremali di una funzione $f : E \rightarrow \mathbf{R}$ a valori reali su un insieme: estremo superiore $\sup f =: \sup \text{Im } f$, estremo inferiore $\inf f =: \inf \text{Im } f$.

Osservazione: i valori estremali possono *non essere valori della funzione* sull' insieme stesso.

Successioni massimizzanti e minimizzanti una successione $x : \mathbf{N} \rightarrow E$ si dice *massimizzante* [*minimizzante*] per f su E se $f(x_n) \xrightarrow{n \rightarrow \infty} \sup_E f$ [$f(x_n) \xrightarrow{n \rightarrow \infty} \inf_E f$].

Osservazione: è fuorviante in generale pensare una successione "estremizzante" come successione convergente. Si nota che tale nozione è data per un dominio E è considerato come semplice insieme, senza alcuna nozione di convergenza in e .

Valore di massimo e valore di minimo $v = \sup_E f$ [$v = \inf_E f$] si dice

valore *massimo* [*minimo*] di f su E se

vi è $p \in E$ per cui $f(p) = \sup_E f$ [$f(p) = \inf_E f$]; nel caso si scriverà $\max_E f$ [$\min_E f$].

Osservazione: il valore di massimo [*minimo*] di una funzione su un insieme è unico.

Punti di massimo, punti di minimo: punti estremali i punti per cui ciò avviene sono detti *punti di massimo* [*di minimo*].

Osservazione: se p è un punto di massimo [*minimo*] la successione costante $x_n = p$, $n \in \mathbf{N}$ è una particolare successione massimizzante [*minimizzante*].

Osservazione: i punti di massimo [*minimo*] possono essere più d' uno.

Studio locale

Punti di massimo relativo, punti di minimo relativo $p \in M$ è un punto di *massimo* [*minimo*] relativo di f su (M, d) se

vi è $U \subseteq M$ intorno in (M, d) di p per cui $f(p) = \max_U f$ rispettivamente [$f(p) = \min_U f$].

Si dicono *stretti* se $f(p) \neq f(x)$ $x \in U \setminus \{p\}$.

I°.2 : OTTIMIZZAZIONE : CRITERI DI ESISTENZA

Weierstrass astratto Se $f : (M, d) \rightarrow \mathbf{R}$ è *continua* e (M, d) *compatto*, allora

f assume massimo su M e assume minimo su M .

Osservazione: in particolare f è *limitata* su M .

Per il teorema di Bolzano Weierstrass (i sottoinsiemi di \mathbf{R}^N compatti per la convergenza euclidea sono tutti e soli i sottoinsiemi chiusi e limitati per la distanza euclidea) si ha:

Weierstrass Se $f : E \subseteq \mathbf{R}^N \rightarrow \mathbf{R}$ è continua lungo E , $E = \bar{E} \subseteq B(0_{\mathbf{R}^N}, r)$ per qualche $r > 0$ allora: f assume massimo su E e assume minimo su E .

Corollario: se $f : A \subseteq \mathbf{R}^N \rightarrow \mathbf{R}$ è semicontinua inferiormente lungo $A = A^\circ$ aperto di \mathbf{R}^N , e vi è $p \in A$ per cui

$$\liminf_{x \in A, |x|_{\mathbf{R}^N} \rightarrow +\infty} f(x) \geq f(p), \quad \liminf_{x \in A, \text{dist}(x, \partial A) \rightarrow 0} f(x) \geq f(p), \quad \text{allora } f \text{ assume minimo su } E.$$

II° : OTTIMIZZAZIONE : METODI DIRETTI

- Semplice uso di disequaglianze e valutazione.

Esempio: $f(x, y, z, w) = x^2 + y^2 + z^2 + w^2$, $E = \{(x, y, z, w) : xyzw = 1\}$.

- $f(n, \frac{1}{n}, 1, 1) = n + \frac{1}{n^2} + 2 \xrightarrow{n \rightarrow \infty} +\infty$, $(n, \frac{1}{n}, 1, 1) \in E$, quindi: $\sup_E f = +\infty$.

- $f(x, y, z, w)$ [Cauchy-Scharz] $\geq 2\sqrt{x^2 + y^2}\sqrt{z^2 + w^2}$ [C.-S.] $\geq 2\sqrt{2xy}\sqrt{2zw}$ su $E = 4$, $f(1, 1, 1, 1) = 4$, $(1, 1, 1, 1) \in E$, quindi: $\inf_E f = \min_E f = 4$.

- Per il problema di massimo [*minimo*]: uso dei teoremi di esistenza, uso di algoritmi che ad ogni passo aumentano [*diminuiscono*] il valore della funzione, riduzione dei candidati punti

estremali, valutazione ed eventuale *confronto* dei valori delle funzione sui diversi sottoinsiemi a cui ci si è ristretti.

Esempio: $f(p, q, r)$ = area del triangolo $T(p, q, r)$ nel piano, di vertici p, q ed r su una circonferenza unitaria: $E = S \times S \times S \subseteq \mathbf{R}^6$, S = circonferenza unitaria di centro c .

- - $f(p, p, p) = 0, p \in S, f \geq 0$, quindi: $\min_E f = 0$.

- - f è continua rispetto alle coordinate in \mathbf{R}^6 , E è compatto, quindi: esiste $\max_E f$:

i- se $c \notin T(p, q, r)$, con $q \in S$ intermedio tra $p, r \in S$, e $q' \in S$ è il simmetrico di q rispetto a c , allora $f(p, q', r) > f(p, q, r)$ e $c \in T(p, q', r)$:

quindi: $\max_E f = \max\{f(p, q, r) : (p, q, r) \in E, c \in T(p, q, r)\}$.

ii- se $c \in T(p, q, r)$, con $q \in S$ intermedio tra $p, r \in S$, $dist(p, q) \neq dist(q, r)$, e $q' \in S$ è il punto di intersezione tra S e il semiasse perpendicolare al segmento $[p; r]$ per c allora $f(p, q', r) > f(p, q, r)$ e $dist(p, q') = dist(q', r)$: quindi esistendo il massimo è assunto sui triangoli equilateri:

$\max_E f = \max\{f(p, q, r) : (p, q, r) \in E, c \in T(p, q, r), T(p, q, r) \text{ equilatero}\}$.

iii- se $(p, q, r) \in E$ con $T(p, q, r)$ equilatero allora: $\max_E f = f(p, q, r) = 3 \frac{\sqrt{3}}{4}$.

Osservazione: nel secondo esempio ci si riduce ad una solo insieme di candidati punti di massimo (i triangoli equilateri inscritti) con un "numero finito di riduzioni". Inoltre su tale insieme la funzione è costante.

Le mosse di riduzione potrebbero essere invece infinite dando origine a successione massimizzante $x_n, n \in \mathbf{N}$.

Andrebbero quindi confrontati tutti i limiti dei valori di f su tali successioni: $\lim_{n \rightarrow \infty} f(x_n)$. Il massimo tra questi è il valore massimo di f .

Avendo l'esistenza a priori si può non badare all'esistenza del limite delle successioni $(x_n)_{n \in \mathbf{N}}$.

III° : OTTIMIZZAZIONE : METODI INDIRETTI

Protocollo generale:

1) criteri di esistenza

2.1) elenco punti degli elementi del dominio che soddisfano le condizioni necessarie,

2.2) elenco degli elementi del dominio per cui non ha senso verificare tali condizioni necessarie

3) valutazione della funzione negli elementi dei precedenti "elenchi":

il massimo [minimo] tra questi valori sarà il massimo [minimo] su tutto il dominio (avendo l'esistenza del valore massimo [minimo] della funzione).

A) Condizioni differenziali necessarie per punti interni

A.1 condizioni al primo ordine - Se $f : E \subseteq \mathbf{R}^N \rightarrow \mathbf{R}$, $p \in E^o$ punto interno di massimo

o minimo relativo di f , e per $v \in \mathbf{R}^N \setminus \{0_{\mathbf{R}^N}\}$ esiste $\frac{\partial f}{\partial v}(p)$, allora: $\frac{\partial f}{\partial v}(p) = 0$.

Dimostrazione: posto $g(t) = f(p+tv)$, $|t| \leq \frac{1}{|v|_N} dist(p, \partial E)$, si ha che $t = 0$ è punto di massimo

[o minimo] relativo di g interno al segmento ove è definita: $\begin{cases} \frac{f(p+tv)-f(p)}{t} \leq 0, & t > 0 \\ \frac{f(p+tv)-f(p)}{t} \geq 0, & t < 0 \end{cases}$.

Pertanto $0 = g'(0) = \frac{\partial f}{\partial v}(p)$.

- Quindi se f è differenziabile si ha $\nabla f(p) = 0_{\mathbf{R}^N}$.

Punti stazionari e valori stazionari: - se $f : E \subseteq \mathbf{R}^N \rightarrow \mathbf{R}^m$, $p \in E$ si dice *stazionario* (per $m = 1$ è detto anche *critico*) se f è differenziabile in p e $D_p f = 0_{m \times N}$;

- se p è un punto stazionario di f il valore $f(p)$ si dice *valore stazionario* ($m = 1$ critico).

A.2 condizioni al secondo ordine - Se $f : E \subseteq \mathbf{R}^N \rightarrow \mathbf{R}$, $p \in E^\circ$ punto interno di massimo

[minimo] relativo di f e per $v \in \mathbf{R}^N \setminus \{0_{\mathbf{R}^N}\}$ esiste $\frac{\partial^2 f}{\partial v^2}(p)$, allora: $\frac{\partial^2 f}{\partial v^2}(p) \leq 0 [\geq 0]$.

Dimostrazione: posto $g(t) = f(p+tv)$, $|t| \leq \frac{1}{|v|_N} \text{dist}(p, \partial E)$, si ha che $t = 0$ è punto di massimo [minimo] relativo di g interno al segmento ove è definita. Considerando lo sviluppo di Taylor al secondo ordine della funzione di una variabile g si ha:

$$0 \geq [\leq] f(p+tv) - f(p) = \frac{t^2}{2} \frac{\partial^2 f}{\partial v^2}(p) + o(t^2).$$

Pertanto dividendo per t^2 e passando al limite per $t \rightarrow 0$: $0 \geq [\leq] \frac{\partial^2 f}{\partial v^2}(p)$.

- Quindi se f è $C^2(E^\circ)$ si ha che la matrice Hessiana $Hf(p)$ di f in p è semidefinita positiva [semidefinita negativa]: ${}^t v Hf(p) v = \frac{\partial^2 f}{\partial v^2}(p) \leq 0 [\geq 0]$.

B) Condizioni differenziali necessarie per punti interni di varietà

B.1 condizioni al primo ordine su sottovarietà

B.1.1 condizioni al primo ordine - Se $f : E \subseteq \mathbf{R}^N \rightarrow \mathbf{R}$, E sottovarietà M dimensionale C^K , $K \geq 1$, $q \in iE$ punto interno di massimo o minimo relativo di f in E , f definita in un intorno U in \mathbf{R}^N di q e differenziabile q . Allora:

- il gradiente tangenziale di f in q lungo E è nullo: $\nabla^E f(q) = 0_{\mathbf{R}^N}$,
- cioè la proiezione ortogonale di $\nabla f(q)$ sulla giacitura T_q del tangente in q ad E è 0_{T_q} ,
- cioè $\nabla f(q)$ è ortogonale a T_q ,
- cioè, nel caso sia anche $\nabla f(q) \neq 0_{\mathbf{R}^N}$, il livello $L = \{x \in U : f(x) = f(q)\}$ è tangente in q a E : $T_q^L \supseteq T_q^E$.

Dimostrazione: - vi è intorno $W \subseteq U$ aperto in \mathbf{R}^N , un aperto $A \subseteq \mathbf{R}^M$, e una funzione $\Psi : A \rightarrow E \cap W \subseteq \mathbf{R}^N$, $\Psi = \Psi(y)$, $y = (y_1, \dots, y_M)$, bigettiva, continua, con *inversa continua*, $C^K(A)$ con differenziale di rango massimo. Sia $a \in A$ per cui $\Psi(a) = q$.

Per ipotesi di estremalità, a patto di rimpicciolire W , si può inoltre assumere che q sia di massimo [minimo] di f su $E \cap W$.

- Quindi a è un punto di massimo [minimo] di $f \circ \Psi$, quindi $\nabla^y (f \circ \Psi) = 0_{\mathbf{R}^M}$.
- per la regola della catena $0_{\mathbf{R}^M} = \nabla^y (f \circ \Psi) = \nabla^y \Psi(a) \nabla^x f(q)$: cioè $\left\langle \frac{\partial \Psi}{\partial y_i}(a) \cdot \nabla f(q) \right\rangle_{\mathbf{R}^N} = 0$.
- Ma $\frac{\partial \Psi}{\partial y_i}(a)$, $1 \leq i \leq M$ è una base di T_q . Quindi $\nabla f(q)$ è ortogonale a T_q e pertanto la sua proiezione ortogonale su di esso è nulla.

Punti stazionari e valori stazionari: - se $f : E \subseteq \mathbf{R}^N \rightarrow \mathbf{R}^m$, E sottovarietà M dimensionale C^K , $K \geq 1$, $p \in iE$ si dice *stazionario tangenziale* o per $m = 1$ anche *stazionario tangenziale* se f è differenziabile in p e $J^E f(p) = 0_{m \times N}$;

- se p è un punto stazionario di f il valore $f(p)$ si dice *valore stazionario (critico) tangenziale*.

B.1.2 moltiplicatori di Lagrange Spesso una E sottovarietà M dimensionale di \mathbf{R}^N si presenta come luogo di zeri di una funzione $g = (g_1, \dots, g_{N-M})$, con differenziale di rango massimo. Conviene formulare la nozione di punto stazionario tangenziale in termini dei ∇g_j , $1 \leq j \leq N - M$ che sono una base dell'ortogonale della giacitura del tangente ad E in q .

Siano $q \in \mathbf{R}^N$, W suo intorno, $g : W \rightarrow \mathbf{R}^{N-M}$, C^1 , $f : W \rightarrow \mathbf{R}$,
 $g(q) = \bar{h}$ e $E = \{x \in W : g(x) = \bar{h}\}$.

si definisce la funzione *Lagrangiana* $\mathcal{L} : W \times \mathbf{R}^{N-M} \rightarrow \mathbf{R}$:

$$\mathcal{L}(x, \lambda, h) = \mathcal{L}(x, \lambda) = f(x) - \langle (g(x) - h) \cdot \lambda \rangle_{\mathbf{R}^{N-M}}.$$

Se: i- f è differenziabile in q lungo E ,

ii- $Jg(q)$ ha rango massimo $N - M$,

allora sono equivalenti

i'- q è punto stazionario tangenziale di f lungo E ,

ii'- vi è $\bar{\lambda} \in \mathbf{R}^{N-M}$ per cui $(q, \bar{\lambda})$ è punto stazionario di \mathcal{L} nel complesso delle variabili (x, λ) ,

iii' - vi è $\bar{\lambda} \in \mathbf{R}^{N-M}$ per cui $(q, \bar{\lambda})$ è soluzione del sistema di Lagrange:

$$\begin{cases} Jf(x) = J\langle g(x) \cdot \lambda \rangle & x \in W \\ g(x) = \bar{h} & x \in W \end{cases}$$

ovvero

$$SL_{\bar{h}} \begin{cases} \frac{\partial f}{\partial x_1}(x) = \lambda_1 \frac{\partial g_1}{\partial x_1}(x) + \dots + \lambda_{N-M} \frac{\partial g_{N-M}}{\partial x_1}(x) & x \in W \\ \vdots \\ \frac{\partial f}{\partial x_N}(x) = \lambda_1 \frac{\partial g_1}{\partial x_N}(x) + \dots + \lambda_{N-M} \frac{\partial g_{N-M}}{\partial x_N}(x) & x \in W \\ g_1(x) = \bar{h}_1 & x \in W \\ \vdots \\ g_{N-M}(x) = \bar{h}_{N-M} & x \in W \end{cases}$$

cioè

$$\begin{cases} \nabla f(x) = \lambda_1 \nabla g_1(x) + \dots + \lambda_{N-M} \nabla g_{N-M}(x) = \nabla g(x) \lambda & x \in W \\ g(x) = \bar{h} & x \in W \end{cases}$$

Dimostrazione: che ii' sia equivalente a iii' si ha dalla linearità delle derivazioni:

$$\frac{\partial \mathcal{L}}{\partial x_i} = \frac{\partial f}{\partial x_i} - \left\langle \frac{\partial g}{\partial x_i} \cdot \lambda \right\rangle, \quad 1 \leq i \leq n, \quad \text{e} \quad \frac{\partial \mathcal{L}}{\partial \lambda_j} = g_j - h_j, \quad 1 \leq j \leq N - M.$$

Per vedere che i' è equivalente a iii' nella sua ultima formulazione: $\nabla f(q)$ (ipotesi i) è ortogonale al tangente in q ad E (esiste per il teorema delle funzioni implicite per ii) se e solo se è combinazione lineare dei $\nabla g_j(q)$, $1 \leq j \leq N - M$: base (ipotesi ii) dell'ortogonale al tangente.

Osservazione: *significato dei moltiplicatori.* Se inoltre f , g sono funzioni C^2 , e la matrice Hessiana di $\mathcal{L}(x, \lambda, \bar{h})$ è non singolare si ha la seguente interpretazione dei moltiplicatori come *derivate, rispetto ai valori del vincolo, dei valori stazionari tangenziali.*

Infatti $(q, \bar{\lambda})$ risolve il sistema di Lagrange $SL_{\bar{h}}$: è nel luogo di zeri di

$\mathcal{G}(x, \lambda, h) = (\nabla f(x) - \nabla \langle g(x) \cdot \lambda \rangle, g(x) - h) : W \times \mathbf{R}^{N-M} \times \mathbf{R}^{N-M} \rightarrow \mathbf{R}^N \times \mathbf{R}^{N-M}$. In un intorno di $(q, \bar{\lambda}, \bar{h})$ sono soddisfatte le ipotesi del teorema del Dini dal minore dato da

$\frac{\partial \mathcal{G}}{\partial (x_1 \dots x_N \lambda_1 \dots \lambda_{N-M})}$.

Si ottiene che vi è un intorno V di \bar{h} e un intorno U di $(q, \bar{\lambda})$ per cui:

per ogni $h \in V$ vi è un'unica soluzione C^1 , $(x(h), \lambda(h))$ di SL_h con la condizione aggiuntiva $(x, \lambda) \in U$. Inoltre $h \rightarrow (x(h), \lambda(h))$ è C^1 . Derivando rispetto alle h_j , $1 \leq j \leq N - M$, il valore $f(x, \lambda)$, dalla regola della catena e per la condizione $g(x(h)) = h$ in SL_h , per cui $\nabla^h(g \circ x)(h) = Id_{N-M \times N-M}$, si ottiene:

$$\frac{\partial f(x(h))}{\partial h_j} = \lambda_j(h).$$

B.2 condizioni al secondo ordine Siano $q \in \mathbf{R}^N$, W suo intorno, $g : W \rightarrow \mathbf{R}^{N-M}$, C^2 con differenziale di rango massimo in q , $f : W \rightarrow \mathbf{R}$, C^2 , $g(q) = \bar{h}$ e $E = \{x \in W : g(x) = \bar{h}\}$. Se q è punto di minimo [massimo] relativo di f in E , se $\bar{\lambda}$ è un moltiplicatore ad esso associato, detta \mathcal{H}_q la matrice Hessiana in q di $L(x) = \mathcal{L}(x, \bar{\lambda})$, e detta $P = P_q^E$ la proiezione ortogonale sulla giacitura T_q^E (${}^t P = P$, $PP = P$, $\text{Im}P = T_q^E$) dello spazio tangente ad E in q , si ha:

$$P\mathcal{H}_qP \text{ è semidefinita positiva [negativa].}$$

Dimostrazione: sia q punto di minimo relativo. Si tratta di usare, grazie al teorema del Dini, come per la condizione necessaria del primo ordine, una carta locale (Ψ, A) $A \subseteq \mathbf{R}^M$, $\Psi(a) = q$, per descrive un intorno di q sul vincolo E . Quindi imporre le condizioni necessarie del secondo ordine sulla matrice Hessiana di $f \circ \Psi$, considernado che $\nabla f(q) = \nabla g(q)\bar{\lambda}$:

$$\frac{\partial^2 f \circ \Psi}{\partial y_j \partial y_k}(y) = \dots = \left\langle (Hf)(\Psi(y)) \frac{\partial \Psi}{\partial y_k} \cdot \frac{\partial \Psi}{\partial y_j} \right\rangle + \left\langle \frac{\partial^2 \Psi}{\partial y_j \partial y_k} \cdot (\nabla f)(\Psi(y)) \right\rangle$$

valutando per $y = a$:

$$\frac{\partial^2 f \circ \Psi}{\partial y_j \partial y_k}(a) = \left\langle Hf(q) \frac{\partial \Psi}{\partial y_k} \cdot \frac{\partial \Psi}{\partial y_j} \right\rangle + \left\langle \nabla g(q)\bar{\lambda} \cdot \frac{\partial^2 \Psi}{\partial y_j \partial y_k} \right\rangle$$

Derivando due volte il sistema $g(\psi(y)) = \bar{h}$ si ottiene anche

$$\left\langle (Hg_i)(\Psi(y)) \frac{\partial \Psi}{\partial y_k} \cdot \frac{\partial \Psi}{\partial y_j} \right\rangle + \left\langle \nabla(g_i)(\Psi(y)) \cdot \frac{\partial^2 \Psi}{\partial y_j \partial y_k} \right\rangle = 0, \quad 1 \leq i \leq N - M$$

valutando in $y = a$ e sostituendo nella precedente eguaglianza

$$\frac{\partial^2 f \circ \Psi}{\partial y_j \partial y_k}(a) = \left\langle [Hf(q) - H(\langle g \cdot \bar{\lambda} \rangle)(q)] \frac{\partial \Psi}{\partial y_k} \cdot \frac{\partial \Psi}{\partial y_j} \right\rangle = \left\langle \mathcal{H}_q \frac{\partial \Psi}{\partial y_k} \cdot \frac{\partial \Psi}{\partial y_j} \right\rangle$$

essendo $\frac{\partial \Psi}{\partial y_j}$, $1 \leq j \leq N - M$ una base del tangente, ciò comporta che \mathcal{H}_q è semidefinita positiva [semidefinita negativa] sul tangente e quindi per ogni $w \in \mathbf{R}^N$:

$$\langle P\mathcal{H}_qPw \cdot w \rangle \geq 0.$$

C) Condizioni differenziali necessarie per punti di frontiera e bordo

C1.1 Se $f : E \subseteq \mathbf{R}^N \rightarrow \mathbf{R}$, differenziabile in $p \in iF$, $F \subseteq \partial E$ varietà, punto di massimo [minimo] relativo di f , allora q è un punto stazionario tangenziale di f lungo F .

C1.2 Se $f : \bar{A} \subseteq \mathbf{R}^N \rightarrow \mathbf{R}$, A aperto, f differenziabile in $p \in \partial A$, punto C^1 regolare per A e punto di massimo [minimo] relativo di f , allora per ogni $v \in \mathbf{R}^N \setminus \{0_{\mathbf{R}^N}\}$ esterno ad A in p :

$$\frac{\partial f}{\partial v}(p) \geq 0 \text{ [} \leq 0 \text{]}.$$

Dimostrazione: per definizione vi è W intorno di p e una g funzione $C^1(U)$ a valori reali, $\nabla g \neq 0$, per cui:

$$W \cap \partial A = \{x : g(x) = 0\}, \quad W \cap A = \{x \in W : g(x) < 0\}.$$

Ora $\frac{\nabla g(p)}{|\nabla g(p)|} =: \nu$ è la normale esterna ad A in p , quindi $\langle v \cdot \nu \rangle > 0$, $v = a^2\nu + w$ con $a \neq 0$ e w tangente a ∂A in p .

Si ha quindi, essendo p punto critico tangenziale per ∂A , che $\frac{\partial f}{\partial w}(p) = 0$:

$$\frac{\partial f}{\partial v}(p) = a^2 \frac{\partial f}{\partial v}(p) + \frac{\partial f}{\partial w}(p) = a^2 \frac{\partial f}{\partial v}(p). \text{ Basta mostrare che } \frac{\partial f}{\partial v}(p) \geq 0.$$

Ora per $t > 0$ si ha $g(p - t\nu) = g(p) - \frac{\partial g}{\partial \nu}(p)t + o(t) = -|\nabla g(p)|t + o(t)$. Quindi per $t > 0$ abbastanza piccolo $g(p - t\nu) < 0$ cioè $p - t\nu \in A$.

Quindi per $t > 0$ piccolo ha senso calcolare $f(p - t\nu)$ che ha $t = 0$ come punto di massimo locale: quindi

$$\frac{f(p - t\nu) - f(p)}{-t} \geq 0, \text{ passando al limite per } t \rightarrow 0^+ \frac{\partial f}{\partial \nu}(p) \geq 0.$$

Osservazione: - la condizione più forte $\frac{\partial f}{\partial v}(p) > 0$ per ogni v esterno in p , non è sufficiente per avere un massimo relativo: $f(x, y) = y - x^3$, $p = (0, 0)$, $A = \{(x, y) : y < 0\}$, $\nabla f(0, 0) = (0, 1)$, v è esterno ad A in $(0, 0)$ se e solo se $v = (a, b)$ con $b > 0$.

- Chiaramente, come ogni condizione necessaria, C1.2 può essere usata in negativo per mostrare che un punto non è di massimo relativo.

C2 Se E è una sottovarietà con bordo, $p \in bE$, f differenziabile in p , suo punto di massimo [minimo] relativo su E , se v è un vettore tangente a E in p esterno non nullo, allora

$$\frac{\partial f}{\partial v}(p) \geq 0 \quad [\leq 0].$$

Dimostrazione; ci si riconduce a C1.2 con carte locali.

Protocollo

1) ESISTENZA

2.1.1) SI TROVANO I PUNTI STAZIONARI INTERNI DEL DOMINIO

2.1.2) SI TROVANO I PUNTI STAZIONARI TANGENZIALI SU SOTTOINSIEMI DEL DOMINIO CHE SIANO VARIETA'

2.1.3) SI TROVANO I PUNTI OVE NON SONO SODDISFATTE LE IPOTESI DELLE CONDIZIONI NECESSARIE:

- i punti ove la funzione da ottimizzare non è differenziabile
- e.g.: i punti su insiemi di livello di funzioni ivi non differenziabili o con differenziale non di rango massimo

3.1) CALCOLO DEI VALORI DELLA FUNZIONE NEI PUNTI DEI PRECEDENTI ELENCHI

3.2) CONFRONTO DEI VALORI OTTENUTI

- il massimo tra questi è il valore massimo della funzione
- il minimo tra questi è il valore minimo della funzione

Osservazione: per 2.1.2 tre metodi

- o si trova una parametrizzazione regolare esplicita del sottoinsieme che interessa e ci si riconduce a 2.1.1 con meno variabili;
- o si esprime il sottoinsieme che interessa come luogo di zeri regolare e si usa il metodo dei moltiplicatori di Lagrange;
- o si esprime il sottoinsieme che interessa come luogo di zeri regolare e si usa direttamente il teorema delle funzioni implicite riducendosi a 2.1.1 e calcolandole derivate delle funzioni implicite.

IV° : STUDIO LOCALE : CONDIZIONI DIFFERENZIALI SUFFICIENTI

A') condizioni sufficienti per punti della parte interna

A'.1 stretta segnatura dell'Hessiana nel punto Se $f \in C^2(W)$, W intorno di p in \mathbf{R}^N e

- p è un punto critico per f : $\nabla f(p) = 0_{\mathbf{R}^N}$,
- la matrice Hessiana di f in p è *strettamente definita* positiva [negativa]:

$$v \in \mathbf{R}^N \setminus \{0_{\mathbf{R}^N}\} \quad {}^t v Hf(p) v > 0, \quad \sum_{i,j=1}^N \frac{\partial^2 f}{\partial x_i \partial x_j}(p) v_i v_j > 0 \quad [< 0]$$

allora p è un punto di *minimo* [massimo] relativo stretto di f .

Dimostrazione: va studiato il segno di $f(x) - f(p)$ per x abbastanza vicino a p .

- Si usa lo sviluppo di Taylor in più variabile del *secondo ordine* e centro p per f , tenendo presente che p è un punto stazionario:

$$f(x) - f(p) = \frac{1}{2} \langle Hf(p)(x-p) \cdot (x-p) \rangle + o(|x-p|^2), \quad x \rightarrow p$$

$$\frac{f(x) - f(p)}{|x-p|^2} = \frac{1}{2} \left\langle Hf(p) \frac{x-p}{|x-p|} \cdot \frac{x-p}{|x-p|} \right\rangle + o(1), \quad x \rightarrow p$$

- - Il polinomio omogeneo di secondo grado $\langle Hf(p)v \cdot v \rangle$ è una funzione continua, quindi, per il teorema di Weierstrass, assume valore minimo sulla sfera unitaria $S^{N-1}(0_{\mathbf{R}^N}, 1)$ che è un compatto. Per qualche w con $|w|_N = 1$:

$$\mu =: \min_{v:|v|_N=1} \langle Hf(p)v \cdot v \rangle = \langle Hf(p)w \cdot w \rangle > 0.$$

In particolare $\left\langle Hf(p) \frac{x-p}{|x-p|} \cdot \frac{x-p}{|x-p|} \right\rangle \geq \mu > 0$.

- - - Concludendo si considera W intorno di p per cui per $x \in W$ si ha $|o(1)| < \frac{\mu}{4}$: per $x \in W$

$$\frac{f(x) - f(p)}{|x-p|^2} = \frac{1}{2} \left\langle Hf(p) \frac{x-p}{|x-p|} \cdot \frac{x-p}{|x-p|} \right\rangle + o(1) \geq \frac{\mu}{2} - \frac{\mu}{4} > 0.$$

A'.2 segnatura dell'Hessiana nell'intorno Se $f \in C^2(W)$, W intorno di p in \mathbf{R}^N e

- p è un punto critico per f : $\nabla f(p) = 0_{\mathbf{R}^N}$,
- per qualche $r > 0$ per cui $B = B(p, r) \subseteq W$
- la matrice Hessiana di f in ogni $z \in B(p, r)$ è *definita* positiva [negativa]:

$$v \in \mathbf{R}^N \setminus \{0_{\mathbf{R}^N}\} \quad {}^t v Hf(z) v \geq 0, \quad \sum_{i,j=1}^N \frac{\partial^2 f}{\partial x_i \partial x_j}(z) v_i v_j \geq 0 \quad [\leq 0]$$

allora p è un punto di *minimo* [massimo] relativo di f .

Dimostrazione: Come sopra usando piuttosto lo sviluppo di Taylor del *primo ordine con resto di Lagrange* di f in p : per $|x-p| < r$ vi è $\xi \in (p; x)$ per cui

$$\frac{f(x) - f(p)}{|x-p|^2} = \frac{1}{2} \left\langle Hf(\xi) \frac{x-p}{|x-p|} \cdot \frac{x-p}{|x-p|} \right\rangle \geq 0.$$

Osservazione: Essendo $f \in C^2(B)$ il fatto che l'Hessiana sia definita positiva è equivalente al fatto che f è una *funzione convessa* sul convesso B .

È un fatto generale che un punti minimo relativo di una funzione convessa su un convesso è un punto di minimo assoluto.

A'.3 punti di sella Se $f \in C^2(W)$, W intorno di p in \mathbf{R}^N e

- p è un punto critico per f : $\nabla f(p) = 0_{\mathbf{R}^N}$,
- $\max_{v:|v|_N=1} \langle Hf(p)v \cdot v \rangle = M > 0, \quad \min_{v:|v|_N=1} \langle Hf(p)v \cdot v \rangle = \mu < 0,$

allora p non è *ne punto di minimo relativo ne punto di massimo relativo* di f .

Dimostrazione: per il teorema di Weierstrass esistono $W, w \in \mathbf{R}^N$ di norma unitaria per cui:

$$\begin{aligned} \max_{v:|v|_N=1} \langle Hf(p)v \cdot v \rangle &= \langle Hf(p)W \cdot W \rangle = M > 0, \\ \min_{v:|v|_N=1} \langle Hf(p)v \cdot v \rangle &= \langle Hf(p)w \cdot w \rangle = \mu < 0, \end{aligned}$$

Se p fosse di massimo o minimo relativo per f lo sarebbe anche per le sue restrizioni a sottoinsiemi a cui p appartiene. In particolare alle rette $p + tW, t \in \mathbf{R}, p + tw, t \in \mathbf{R}$:

$$\frac{\partial^2 f}{\partial w^2}(p) = \langle Hf(p)w \cdot w \rangle = \mu < 0$$

quindi p non può essere per le condizioni necessarie di minimo relativo,

$$\frac{\partial^2 f}{\partial W^2}(p) = \langle Hf(p)W \cdot W \rangle = M > 0$$

quindi p non può essere per le condizioni necessarie di massimo relativo.

B') condizioni sufficienti per punti stazionari tangenziali interni

B'.1. Siano $q \in \mathbf{R}^N, W$ suo intorno, $g: W \rightarrow \mathbf{R}^{N-M}, C^2$ con differenziale di rango massimo in $q, f: W \rightarrow \mathbf{R}, C^2, g(q) = \bar{h}$ e $E = \{x \in W : g(x) = \bar{h}\}$.

Si consideri la funzione *Lagrangiana* $\mathcal{L}(x, \lambda, h) = \mathcal{L}(x, \lambda) = f(x) - \langle (g(x) - h) \cdot \lambda \rangle_{\mathbf{R}^{N-M}}$.

Se q è punto *stazionario tangenziale* di f lungo E , se $\bar{\lambda}$ è un moltiplicatore ad esso associato, detta \mathcal{H}_q la matrice Hessiana in q di $L(x) = \mathcal{L}(x, \bar{\lambda}, \bar{h})$, e detta $P = P_q^E$ la proiezione ortogonale sulla giacitura T_q^E (${}^t P = P, PP = P, \text{Im} P = T_q^E$) dello spazio tangente ad E in q , se:

$P\mathcal{H}_q P$ ha somma delle molteplicità degli autovalori strettamente positivi [negativi] eguale ad M , cioè la matrice \mathcal{H}_q ristretta alla giacitura del piano tangente ad E in q è strettamente definita positiva [negativa] allora

q è un punto di minimo [massimo] relativo stretto di f su E .

Dimostrazione: come fatto in B.2 si usano carte locali in q punto di minimo relativo e ci si riconduce ad usare A'.1. Si consideri quindi una carta locale (Ψ, A) $A \subseteq \mathbf{R}^M, \Psi(a) = q, \Psi(A) = E \cap W, W$ intorno di $q, D_a \Psi$ di rango massimo.

- $\frac{\partial \Psi}{\partial y_j}(a)$ è una base della giacitura T^E del tangente ad E in q , ed essendo $q = \Psi(a)$ un punto

critico tangenziale $\left\langle \nabla f(q) \cdot \frac{\partial \Psi}{\partial y_j}(a) \right\rangle = 0$.

- - Poichè per ogni $w \perp T_q^E$ si ha $Pw = 0_{\mathbf{R}^N}$ si ha anche $P\mathcal{H}_q Pw = 0_{\mathbf{R}^N}$. Quindi 0 è autovalore con molteplicità geometrica $N - M$ e il suo autospazio è l'ortogonale a T_q^E .

Per ipotesi la matrice simmetrica $P\mathcal{H}_q P$ ha somma delle molteplicità geometriche degli autovalori strettamente positivi eguale ad M , quindi gli autospazi di autovalore positivo danno come somma diretta proprio T_q^E che ha dimensione M , pertanto:

$$\text{se } v \in T_q^E \setminus \{0_{\mathbf{R}^N}\} \text{ si ha } \langle \mathcal{H}_q v \cdot v \rangle = \langle P\mathcal{H}_q P v \cdot v \rangle > 0.$$

Per ogni $y = (y_1, \dots, y_M) \neq 0_{\mathbf{R}^M}$:

$$\sum_{1 \leq j, k \leq M} \left\langle \mathcal{H}_q \frac{\partial \Psi}{\partial y_j}(a) \cdot \frac{\partial \Psi}{\partial y_k}(a) \right\rangle y_j y_k = \left\langle \mathcal{H}_q \sum_{j=1}^M y_j \frac{\partial \Psi}{\partial y_j}(a) \cdot \sum_{k=1}^M y_k \frac{\partial \Psi}{\partial y_k}(a) \right\rangle > 0,$$

--- Come in B.2 è $\frac{\partial^2 f \circ \Psi}{\partial y_j \partial y_k}(a) = \left\langle [Hf(q) - H(\langle g \cdot \bar{\lambda} \rangle)(q)] \frac{\partial \Psi}{\partial y_k} \cdot \frac{\partial \Psi}{\partial y_j} \right\rangle = \left\langle \mathcal{H}_q \frac{\partial \Psi}{\partial y_k} \cdot \frac{\partial \Psi}{\partial y_j} \right\rangle$, ed essendo a un punto critico di $f \circ \Psi, \nabla^y f \circ \Psi(a) = \nabla^y \Psi(a) \nabla^x f(q) = 0_{\mathbf{R}^M}$, per A'.1 a è un punto di minimo relativo stretto di $f \circ \Psi$.

Quindi, avendo f ed $f \circ \Psi$ egual immagine e potendo assumere che l'immagine di un intorno di a sia intorno relativo in E di $\Psi(a) = q$ (Ψ^{-1} continua), q è un punto di minimo relativo stretto di f su E .

OSSERVAZIONE: nella pratica per verificare una tale condizione sufficiente può esser comodo usar la seguente strategia. Essa usa sia i moltiplicatori, che direttamente il teorema delle funzioni implicite usando come parametrizzazione Ψ la *funzione grafico delle funzioni implicite* di cui *basterà* calcolare le derivate prime.

Ci si ferma al penultimo passaggio della dimostrazione. Ciò evita inoltre di porsi il problema della proiezione ortogonale sul tangente.

i- si considera un punto critico tangenziale q e un moltiplicatore ad esso associato $\bar{\lambda}$;

ii- per il teorema del Dini si può assumere che $E \cap W$ sia il grafico di una funzione ϕ di M variabili a valori in $\mathbf{R}^{N-M} \subseteq \mathbf{R}^N$.

Per esempio delle prime M : $\det \left(\frac{\partial g}{\partial x_{M+1}} \dots \frac{\partial g}{\partial x_N} \right) \neq 0$.

Considerando $\Psi(x_1, \dots, x_M) = (x_1, \dots, x_M, \phi_1(x_1, \dots, x_M), \dots, \phi_{N-M}(x_1, \dots, x_M))$, $a = (q_1, \dots, q_M)$

$$\begin{aligned} (Hf \circ \Psi)(a) &= \nabla \Psi(a) [Hf(q) - H(\langle g \cdot \bar{\lambda} \rangle)(q)] J\Psi(a) = \\ &= [Id_{M \times M} \nabla \phi(q_1, \dots, q_M)] [Hf(q) - H(\langle g \cdot \bar{\lambda} \rangle)(q)] \begin{bmatrix} Id_{M \times M} \\ J\phi(q_1, \dots, q_M) \end{bmatrix}; \end{aligned}$$

iii- si tratta infine di stabilire la segnatura degli autovalori di tale matrice: se strettamente positiva [negativa] q è un minimo [massimo] relativo di f su E .

- Tali matrici sono direttamente calcolabili, e in certi casi ciò può esser più agevole che il calcolo della proiezione ortogonale sulla giacitura dello spazio tangente.

- Nel caso generale, in cui le M variabili indipendenti sono un altro blocco $\det (Jg)^{\phi_1 \dots \phi_m} \neq 0$, si ha $a = (q_{\sigma_1}, \dots, q_{\sigma_M})$ piuttosto che (q_1, \dots, q_M) , e quindi si tratta di permutare solo le righe, rispettivamente le colonne, dell'ultimo fattore matriciale e del primo fattore matriciale.

C') condizioni sufficienti per punti di frontiera o bordo

C'.1 Se $f : \bar{A} \subseteq \mathbf{R}^N \rightarrow \mathbf{R}$, A aperto, $f \in C^1$, $p \in \partial A$, punto C^1 regolare per A , per cui $\nabla f(p) \neq \vec{0}$, ed inoltre

i- p è punto di *massimo* [minimo] relativo di f ristretta a ∂A , e quindi $\nabla f(p) \perp T_p(\partial A)$,

ii- $\nabla f(p)$ è un vettore esterno [interno] ad A in p , cioè, essendo ortogonale in p a ∂A , per ogni $v \in \mathbf{R}^N \setminus \{0_{\mathbf{R}^N}\}$ esterno ad A in p : $\frac{\partial f}{\partial v}(p) > 0$ [< 0],

allora

p è punto di *massimo* [minimo] relativo di f .

Dimostrazione: - per definizione di punto regolare di frontiera vi sono un *cilindro* $U \sim B \times I$ aperto in \mathbf{R}^N , $\psi : B \rightarrow I$ funzione reale C^1 , di $N - 1$ tra le coordinate x_1, \dots, x_N di \mathbf{R}^N , B palla aperta di tale sottospazio $N - 1$ dimensionale, per cui $p \in U$, $U \cap \partial A$ e $U \cap A$ siano rispettivamente il *grafico*, e *l'intersezione* di U con il *sottografico* aperto di ψ . Per comodità si suppone che le coordinate indipendenti siano le prime $N - 1$: $A \cap U = \{(y, z) \in U : z < \psi(y)\}$, $\partial A \cap U = \{(y, z) \in U : z = \psi(y)\}$, $p = (q, \psi(q))$, $q \in B$.

- Si ha che $\nabla f(p) = \lambda(-\nabla \psi(q), 1)$ per qualche $\lambda > 0$ essendo normale esterno al grafico di ψ in $p = (q, \psi(q))$: infatti il piano tangente in $p = (q, \psi(q))$ a $\partial A \cap U = \text{Graf}\psi$ ha equazione $-\langle \nabla \psi(q) \cdot (y - q) \rangle + (z - \psi(q)) = 0$.

- Per continuità di f e di ∇f si può rimpicciolire U in modo che p sia punto di massimo di f sul grafico di ψ e ∇f abbia sempre ultima componente positiva.

Dato quindi $x = (y, z) \in U \cap A$ ($z < \psi(y)$) si ha che $g(z) = f(y, z)$ ha derivata in z eguale a $\frac{\partial f}{\partial z}(y, z) = \langle \nabla f(y, z) \cdot (\vec{0}, 1) \rangle > 0$. Quindi $f(x) = g(z) \leq g(\psi(y)) = f(y, \psi(y)) \leq f(p)$.

Usando parametrizzazioni regolari locali si ottiene quindi

C'.2 Se E è una sottovarietà con bordo C^1 , $p \in bE$, $f \in C^1$, $\nabla^E f(p)$ non nullo, e

i- p è punto di *massimo* [minimo] relativo di f ristretta a bE ,

ii- per ogni $v \in \mathbf{R}^N \setminus \{0_{\mathbf{R}^N}\}$ tangente esterno ad E in p : $\frac{\partial f}{\partial v}(p) > 0$ [< 0],

allora

p è punto di *massimo* [minimo] relativo di f .

[B] per V.Barutello et al. Analisi mat. vol. 2;

[F] per N.Fusco et al. An.Mat. due;

[FS] per N.Fusco et al. Elem. di An. Mat. due, versione semplificata.

[FS] pagg.73-81 (massimi e mininmi relativi in due variabili), pagg. 287-293 (moltiplicatori di Lagrange in due variabili), pagg 293-296 (moltiplicatori in piu' dimensioni);

[B] pagg. 301-307, 310-311 (in due variabili), pagg. 363-365 (condizioni differenziali necessarie e sufficienti per punti interni in piu' variabili), pagg. 369-371 (moltiplicatori di Lagrange);

[F] pagg. 169-176 (condizioni differenziali necessarie e condizioni sufficienti in aperti), pagg 623-634 (moltiplicatori di Lagrange).