ELEMENTI DI CALCOLO DIFFERENZIALE. PARTE II

FAUSTO FERRARI

Materiale propedeutico alle lezioni di Analisi Matematica per i corsi di Laurea in Ingegneria Chimica e per l'Ambiente e il Territorio dell'Università di Bologna. Anno Accademico 2006/2007

1. Alcune nozioni di geometria lineare

Sia v un vettore di $\mathbb{R}^n$. L'insieme

$$\{y \in \mathbb{R}^n : \ y = tv, \ t \in \mathbb{R}\}$$

è la retta di versore v passante per l'origine (forma parametrica).

Analogamente, dato un punto x_0 e un vettore $v \in \mathbb{R}^n$ l'insieme seguente

$$\{y \in \mathbb{R}^n : \ y = x_0 + tv, \ t \in \mathbb{R}\}$$

è la retta passante per x_0 di versore v.

Per ogni $x_1, x_2 \in \mathbb{R}^n$ l'insieme

$$\{y \in \mathbb{R}^n : \ y = x_1 + t(x_2 - x_1), \ t \in [0,1]\}$$

è detto segmento di estremi x_1 e x_2. L'equazione del segmento in forma parametrica si scrive anche nel seguente modo

$$tx_2 + (1-t)x_1,$$

$t \in [0,1]$. Il segmento viene indicato con il simbolo $[x_1, x_2]$.

Richiamiamo sommariamente alcune nozioni di geometria lineare.

Se $L : \mathbb{R}^n \to \mathbb{R}^n$ è un'applicazione lineare rispetto alla base canonica, allora esiste una matrice di numeri M, $n \times n$, tale che per ogni h in $\mathbb{R}^n$

$$L(h) = M \cdot h.$$

Data un'applicazione lineare $L : \mathbb{R}^n \to \mathbb{R}^n$ diremo che un vettore $v \in \mathbb{R}^n$ è un autovettore di autovalore $\lambda \in \mathbb{R}$ se

$$L(v) = \lambda v.$$

Se M è la matrice associata all'applicazione lineare L, allora gli autovalori di L si calcolano risolvendo l'equazione caratteristica

$$\det(M - \lambda I) = 0$$

in $\mathbb{C}$.

Data una matrice di numeri Q, $n \times n$. L'applicazione $P : \mathbb{R}^n \to \mathbb{R}$ così definita, rispetto alla base canonica,

$$P(h) = \langle Qh, h\rangle$$

è detta forma quadratica. Una forma quadratica è definita positiva se per ogni $h \in \mathbb{R}^n$, $h \neq 0$, $P(h) > 0$. In tal caso diremo anche che la matrice Q è definita positiva.

Analogamente una forma quadratica $P = \langle Qh, h\rangle$ è definita negativa se per ogni $h \in \mathbb{R}^n$, $h \neq 0$, $P(h) < 0$ ovvero diremo anche che la matrice Q associata a tale forma quadratica è definita negativa.

Se le precedenti disuguaglianze sono deboli si parlerà di forme quadratiche (e matrici) semidefinite.

Date: 15/02/2004.

Infine se esistono h_1 e h_2 vettori tali che $Q(h_1) < 0$ e $Q(h_2) > 0$, allora diremo che la forma quadratica Q non è definita (o anche che è indefinita).

Vale il seguente risultato. Data una matrice di numeri reali H, $n \times n$, simmetrica. Allora la forma quadratica ad essa associata

$$Q(h) = \langle H \cdot h, h \rangle$$

soddisfa le seguenti disuguaglianze: per ogni $h \in \mathbb{R}^n$

$$\lambda ||h||^2 \leq \langle H \cdot h, h \rangle \leq \Lambda ||h||^2,$$

dove

$$\lambda = \min\{\gamma : \gamma \text{ è autovalore di } H\}$$

$$\Lambda = \max\{\gamma : \gamma \text{ è autovalore di } H\}.$$

Inoltre se v_m è un autovettore associato a λ e v_M è un autovettore associato a Λ, risulta

$$Q(v_m) = \lambda ||v_m||^2 \tag{1}$$

e

$$Q(v_m) = \Lambda ||v_M||^2. \tag{2}$$

È possibile stabilire se una matrice simmetrica è definita positiva o negativa o indefinita in base alla seguente regola. Se H è la matrice simmetrica di ordine $n \times n$, indichiamo con $\mid H_i \mid$ i minori principali, vale a dire i determinanti delle matrici quadrate di ordine inferiore rispetto alla diagonale principale. Se per ogni $i = 1, 2 \cdots, n$, $\mid H_i \mid > 0$, allora la matrice è definita positiva. Se per ogni $i = 1, 2 \cdots, n$, $(-1)^i \mid H_i \mid > 0$, allora la matrice è definita negativa.

Se esiste i pari tale che $\mid H_i \mid < 0$ allora la forma quadratica e la matrice associata è indefinita (cioè non è né definita pos. né def. neg.)

Vediamo un caso. Sia

$$M = \begin{bmatrix} 5, & 3, & -2 \\ 3, & 4, & -1 \\ -2, & -1, & 1 \end{bmatrix}.$$

Allora $M_1 = 5$, $M_2 = 11$, $M_3 = 2$, quindi è definita positiva.

Vale il seguente risultato in $\mathbb{R}^2$ Sia $Q : \mathbb{R}^2 \to \mathbb{R}$ una forma quadratica simmetrica e M la matrice (simmetrica) 2 che caratterizza Q.

1) Se det$M > 0$ e $m_{11} \neq 0$, allora M è definita (positiva o negativa). In particolare
   1a) Se $m_{11} > 0$ e det$M > 0$, allora M è definita positiva.
   1b) Se $m_{11} < 0$ e det$M > 0$, allora M è definita negativa.
2) Se det$M < 0$, allora è indefinita.
3) Se det$M = 0$ e $m_{11} \neq 0$ allora la matrice è semidefinita.

   Le precedenti affermazioni possono essere analogamente riformulate nel caso in cui $m_{11} = 0$ e $m_{22} \neq 0$.

2. Derivate parziali di ordine superiore

Analogamente a quanto accade nel caso di funzioni di una variabile, possiamo definire le derivate parziali di ordine superiore al primo.

Definizione 2.1. *Sia $\Omega \subset \mathbb{R}^2$ e $f : \Omega \to \mathbb{R}$. Supponiamo che esista la derivata parziale f rispetto a x_j, con $j \in \{1, 2\}$ per ogni $x \in I_{k,\epsilon} = \{y \in \mathbb{R}^2 : \ y = x_0 + te_k \ \ -\epsilon < t < \epsilon\}$, $k \in \{1, 2\}$ dove $x_0 \in int(\Omega)$, e_k è il corrispondente vettore della base canonica alla coordinata x_k e ϵ è un* *numero positivo tale che $I_{k,\epsilon} \subset int(\Omega)$. Diremo che f è dotata di derivata parziale di ordine due rispetto a x_j e x_k se esiste*

$$\lim_{t\to 0} \frac{\frac{\partial f}{\partial x_j}(x_0+te_k) - \frac{\partial f}{\partial x_j}(x_0)}{t} = \alpha_{k,j} \in \mathbb{R}.$$

In tal caso scriveremo

$$\frac{\partial^2 f}{\partial x_k \partial x_j}(x_0).$$

per indicare il numero $\alpha_{k,j}$.

Utilizzeremo la segunete simbologia. Sia $\Omega \subset \mathbb{R}^2$ aperto.

$$\begin{aligned} C^2(\Omega) &= \{f:\Omega\to\mathbb{R}:\ f \text{ è dotata di tutte le derivate parziali fino all'ordine} \\ 2 &\quad \text{in ogni punto di } \Omega \text{ e queste sono continue in } \Omega\}. \end{aligned}$$

Si definiscono in modo analogo le derivate parziali di ordine superiore a due e le classi di funzioni C^k, con $k > 2$.

Senza ulteriori ipotesi sulla funzione f, l'ordine con il quale viene calcolata la derivata parziale succesiva è rilevante. Infatti se $j \neq k$, allora

$$\frac{\partial^2 f}{\partial x_k \partial x_j}(x_0) \neq \frac{\partial^2 f}{\partial x_j \partial x_k}(x_0).$$

Tuttavia vale il seguente Teorema che formuliamo in $\Omega \subset \mathbb{R}^n$

Teorema 2.1. *(Schwarz) Se $f \in C^2(\Omega)$ e $\Omega \subset \mathbb{R}^n$ è aperto , allora per ogni $j,k \in \{1,\dots,n\}$ e per ogni $x_0 \in \Omega$*

$$\frac{\partial^2 f}{\partial x_k \partial x_j}(x_0) = \frac{\partial^2 f}{\partial x_j \partial x_k}(x_0).$$

Definizione 2.2. *Sia $x_0 \in \Omega \subset \mathbb{R}^2$, con Ω aperto. Se esistono le derivate parziali di ordine 2 in x_0, allora chiamiamo matrice Hessiana di f in x_0 la matrice quadrata 2×2*

$$Hf(x_0) = \left(\frac{\partial^2 f}{\partial x_k \partial x_j}(x_0)\right)_{1\le k,j\le 2}.$$

Analogamente se $x_0 \in \Omega \subset \mathbb{R}^2$, con Ω aperto. La matrice $n \times n$

$$Hf(x_0) = \left(\frac{\partial^2 f}{\partial x_k \partial x_j}(x_0)\right)_{1\le k,j\le n}$$

è detta matrice Hessiana di f in x_0.

Sia $f \in C^2(\Omega)$, $\Omega \subset \mathbb{R}^2$, allora la matrice Hessiana di f in $x_0 \in \Omega$,

$$Hf(x_0) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(x_0), & \frac{\partial^2 f}{\partial x_2 \partial x_1}(x_0) \\ \frac{\partial^2 f}{\partial x_1 \partial x_2}(x_0), & \frac{\partial^2 f}{\partial x_2^2}(x_0), \end{bmatrix}$$

è simmetrica. Analogamente se $f \in C^2(\Omega)$, $\Omega \subset \mathbb{R}^n$, allora

$$Hf(x_0) = \left(\frac{\partial^2 f}{\partial x_k \partial x_j}(x_0)\right)_{1\le k,j\le n}$$

è simmetrica.

Prima di scrivere la formula di Taylor all'ordine due vediamo l'equivalente del Teorema del valor medio per f.

Teorema 2.2. *Sia $f \in C^1(\Omega)$, con $\Omega \subset \mathbb{R}^2$ e $[x_0, x] \subset int(\Omega) = \Omega$. Allora esiste $\bar{x} \in]x_0, x[$ tale che*

$$f(x) - f(x_0) = \langle \nabla f(\bar{x}), x - x_0 \rangle.$$

Sia $t \to g(t) = f(x_0 + t(x - x_0))$ definita su $[0, 1]$. Notiamo che $g \in C([0, 1])$ e g è derivabile in $]0, 1[$. Dal Teorema del valor medio di Lagrange risulta che esiste $\theta \in]0, 1[$ tale che

$$g(1) - g(0) = g'(\theta).$$

Ciò si traduce nella seguente uguaglianza

$$f(x) - f(x_0) = \langle \nabla f(x_0 + \theta(x - x_0)), x - x_0 \rangle,$$

quindi $\bar{x} = x_0 + \theta(x - x_0)$.

Il Teorema del valor medio di Lagrange è la formula di Taylor con resto di Lagrange al primo ordine. Vediamo che cosa accade se ci arrestiamo al secondo ordine.

Teorema 2.3. *(Formula di Taylor) Sia $f \in C^2(\Omega)$, $\Omega \subset \mathbb{R}^2$. Supponiamo che $[x_0, x] \subset \Omega$. Allora*

$$f(x) = f(x_0) + \langle \nabla f(x_0), x - x_0 \rangle + \frac{1}{2}\langle Hf(x_0)(x - x_0), (x - x_0) \rangle + o(||x - x_0||^2), \;\; x \to x_0.$$

Proof. Forniamo una dimostrazione di questo risultato per $f \in C^3$. Definiamo la funzione ausiliaria $g : [0, 1+\epsilon[\to \mathbb{R}$, $g(t) = f(x_0 + t(x - x_0))$, con ϵ numero positivo eventualmente piccolo in modo che $\{x_0 + t(x - x_0) : \; t \in [0, 1 + \epsilon[\} \subset \Omega$, in quanto per ipotesi $[x_0, x] \subset \Omega$, con Ω aperto . La funzione g è di classe $C^3(] - \epsilon, \epsilon[)$. Quindi per la formula di Taylor con resto di Lagrange risulta

$$g(t) = g(0) + g'(0)t + \frac{1}{2}g''(0)t^2 + \frac{1}{6}g'''(\xi)t^3,$$

con $\xi]0, 1 + \epsilon[$. In particolare risulterà

$$g(1) = g(0) + g'(0) + \frac{1}{2}g''(0) + \frac{1}{6}g'''(\xi),$$

$\xi]0, 1[$. D'altra parte, $g(0) = f(x_0)$ e $g(1) = f(x)$, mentre dal teorema di differenziabilità delle funzioni composte risulta

$$g'(t) = \langle \nabla f(x_0 + t(x - x_0)), x - x_0 \rangle = \sum_{i=1}^{2} \frac{\partial f}{\partial x_i}(x_0 + t(x - x_0))(x_i - x_{0i}),$$

$$g''(t) = (g'(t))' = \sum_{i=1}^{2}\sum_{k=1}^{2} \frac{\partial^2 f}{\partial x_k \partial x_i}(x_0 + t(x - x_0))(x_i - x_{0i})(x_k - x_{0k})$$

e

$$g'''(t) = (g''(t))' = \sum_{j=1}^{2}\sum_{i=1}^{2}\sum_{k=1}^{2} \frac{\partial^3 f}{\partial x_i \partial x_k \partial x_j}(x_0 + t(x - x_0))(x_i - x_{0i})(x_k - x_{0k})(x_j - x_{0j}).$$

Quindi $g'(0) = \langle \nabla f(x_0), x - x_0 \rangle$

$$g''(0) = \sum_{i=1}^{2}\sum_{k=1}^{2} \frac{\partial^2 f}{\partial x_k \partial x_i}(x_0)(x_i - x_{0i})(x_k - x_{0k}) = \langle H(f(x_0))(x - x_0), (x - x_0) \rangle$$

e

$$g'''(\xi) = \sum_{j=1}^{2}\sum_{i=1}^{2}\sum_{k=1}^{2} \frac{\partial^3 f}{\partial x_i \partial x_k \partial x_j}(x_0 + \xi(x - x_0))(x_i - x_{0i})(x_k - x_{0k})(x_j - x_{0j}).$$

Quindi la formula di Taylor vale, perché $g'''(\xi) = o(||x - x_0||^2)$, per $x \to x_0$. □

Ovviamente il risultato vale anche per $n > 2$ nelle stesse ipotesi di regolarità del Teorema precedente. In particolare varrà

$$f(x) = f(x_0) + \langle \nabla f(x_0), x - x_0 \rangle + \frac{1}{2} \langle Hf(x_0)(x - x_0), (x - x_0) \rangle + o(||x - x_0||^2), \;\; x \to x_0,$$

dove $\nabla f(x_0)$ è un vettore $1 \times n$ e $Hf(x_0)$ è una matrice $n \times n$.

3. Classificazione dei punti estremanti

Ricordiamo le seguenti definizioni

Definizione 3.1. *(Punto di max locale) Sia $\Omega \subset \mathbb{R}^n$, $x_0 \in \Omega \cap D(\Omega)$ e $f : \Omega \to \mathbb{R}$. Diremo che x_0 è un punto di massimo locale per f se esiste $r > 0$ tale che per ogni $x \in B(x_0, r) \cap \Omega$,*

$$f(x) \leq f(x_0).$$

Definizione 3.2. *(Punto di min locale) Sia $\Omega \subset \mathbb{R}^n$, $x_0 \in \Omega \cap D(\Omega)$ e $f : \Omega \to \mathbb{R}$. Diremo che x_0 è un punto di minimo locale per f se esiste $r > 0$ tale che per ogni $x \in B(x_0, r) \cap \Omega$,*

$$f(x) \geq f(x_0).$$

La richiesta che x_0 sia anche punto di accumulazione serve per evitare di considerare alcuni casi evidentemente privi di significato, come una funzione definita su $\mathbb{Z}^n$ per la quale tutti i punti risulterebbero di massimo e minimo locale.

Si parlerà di punto di massimo relativo stretto per f qualora x_0 sia punto di massimo relativo per f e

$$f(x) < f(x_0), \;\; \text{per ogni} \;\; x \in B(x_0, r) \cap (\Omega \setminus \{x_0\}).$$

Analoga definizione viene data per il caso di un punto di minimo relativo stretto.

I punti di minimo o massimo locale per una funzione f sono genericamente detti punti estremanti per f.

Per meglio capire la natura del test per mezzo del quale è possibile determinare, qualora esistano, i punti estremanti per una funzione f definita su $\mathbb{R}^2$ o $\mathbb{R}^n$ trattiamo il caso in $\mathbb{R}$ riassunto nel segunete teorema.

Teorema 3.1. *Sia $I \subset \mathbb{R}$ un intervallo $x_0 \in int(I)$ e $f : I \to \mathbb{R}$. Supponimo che esistano $f'(x_0)$ e $f''(x_0)$ reali. Se*

1) $f'(x_0) = 0$
2) $f''(x_0) \neq 0$ *Allora x_0 è un punto estremante relativo forte per f.*
   In particolare se $f'(x_0) = 0$ e $f''(x_0) > 0$, allora x_0 è punto di minimo relativo forte per f, mentre se $f'(x_0) = 0$ e $f''(x_0) < 0$, allora x_0 è punto di massimo relativo forte per f.

Vediamo, nel caso in cui la funzione sia di classe $C^2(I)$ con I aperto, la dimostrazione di questo risultato. Consideriamo la formula di Taylor con resto di Peano nel punto x_0. Allora

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2} f''(x_0)(x - x_0)^2 + o(x - x_0)^2, \;\; x \to x_0$$

e $\xi \in]x_0, x[$, se $x_0 < x$, altrimenti $\xi \in]x, x_0[$. A questo punto ricordiamo che $f'(x_0) = 0$. Allora

$$f(x) - f(x_0) = \frac{1}{2} f''(x_0)(x - x_0)^2 + o(x - x_0)^2, \;\; x \to x_0$$

da cui

$$f(x) - f(x_0) = (x - x_0)^2 (\frac{1}{2} f''(x_0) + o(1)), \;\; x \to x_0. \tag{3}$$

Supponiamo $f''(x_0) > 0$, allora dal teorema della permanenza del segno segue che esiste un intorno di x_0, I_{x_0}, tale che per ogni $x \in I_{x_0} \setminus \{x_0\}$, $(\frac{1}{2} f''(x_0) + o(1)) > 0$. Quindi risulterà positivo anche il membro di destra di (3) cioè $(x-x_0)^2(\frac{1}{2}f''(x_0)+o(1))>0$ per ogni $x\in I_{x_0}\setminus\{x_0\}$ e pertanto anche il membro di sinistra di (3) sarà positivo per ogni $x\in I_{x_0}\setminus\{x_0\}$, ovvero per ogni $x\in I_{x_0}\setminus\{x_0\}$

$$f(x_0)<f(x),$$

cioè x_0 è un punto di minimo relativo forte per f.

Nel caso di funzioni a più variabili il risultato che enunceremo sarà l'analogo del precedente Teorema in una sola variabile. Il ruolo della derivata prima in x_0 sarà preso dal gradiente di f in x_0, mentre quello della derivata seconda in x_0 sarà preso dalla matrice Hessiana di f in x_0. Vediamo ora che l'annullarsi del gradiente di f in un punto interno è condizione necessaria affinché il punto sia estremante. Ovviamente non sarà condizione sufficiente.

Definizione 3.3. *Sia $x_0\in Int(\Omega)$, dove $\Omega\subset\mathbb{R}^2$. Se in x_0 esistono tutte le derivate parziali di ordine uno in x_0 e $\nabla f(x_0)=0$, allora diremo che x_0 è critico.*

Esistono punti che sono critici senza essere punti estremanti, si pensi alla funzione $f:\mathbb{R}^2\to\mathbb{R}$, $f(x,y)=xy$.

$$\nabla f(x,y)=(y,x).$$

L'unico punto critico è $(0,0)$, ma $(0,0)$ non è estremante, perché $f(t,t)>0$, se $t>0$, mentre $f(t,-t)<0$ se $t>0$ e $f(0,0)=0$.

Teorema 3.2. *(Condizione necessaria per l'esistenza di punti estremanti).*

Sia $x_0\in Int(\Omega)$, dove $\Omega\subset\mathbb{R}^2$. Se x_0 ̀punto estremante relativo per $f:\Omega\to\mathbb{R}$ e f è dotata di tutte le derivate partziali in x_0, allora

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

Per la dimostrazione sarà sufficiente considerare per ogni $j\in\{1,2\}$ $g_j:I_j\to\mathbb{R}$ intorno di 0 in $\mathbb{R}$, con I_j eventualmente piccolo, così definita $g_j(t)=f(x_0+te_j)$. Le funzioni g_j sono dotate di derivata in 0 e queste derivate non sono altro che le derivate parziali di f in x_0. Ovvero

$$g_1'(0)=\frac{\partial f}{\partial x_1}(x_0)$$

$$g_2'(0)=\frac{\partial f}{\partial x_2}(x_0).$$

D'altra parte per ogni $j\in\{1,2\}$ si ha che 0 è punto estremante per g_j, della stessa natura di f. Vale a dire: se x_0 è punto di massimo relativo per f, allora anche 0 è punto di massimo relativo per g_j (analogamente per il minimo). Poiché 0 è punto interno per I_j possiamo applicare il Teorema di Fermat a g_j e concludere che

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

Sia $f:\Omega\subset\mathbb{R}^2\to\mathbb{R}$ dotata di tutte le derivate parziali di ordine uno in $x_0\in\mathrm{Int}(\Omega)$ Diremo che x_0 è un punto di sella per f se $\nabla f(x_0)=0$, ma x_0 non è punto estremante per f.

Teorema 3.3. *Sia $x_0\in Int(\Omega)$, $\Omega\subset\mathbb{R}^2$, $f\in C^2(Int(\Omega))$. Se*

1) $\nabla f(x_0)=0$ *(x_0 è critico)*
   e
2) $Hf(x_0)$ *è definita (positiva o negativa).*
   Allora x_0 è punto estremante relativo forte per f. In particolare se $\nabla f(x_0)=0$ e $Hf(x_0)>0$ allora x_0 è un punto di minimo locale forte, mentre se $\nabla f(x_0)=0$ e $Hf(x_0)<0$ allora x_0 è un punto di massimo locale forte.

Inoltre se

1) $\nabla f(x_0)=0$ *(x_0 è critico)*
   e

2) *$Hf(x_0)$ non è definita (né positiva né negativa), allora x_0 è punto di sella.*

La dimostrazione è analoga al caso unidimensionale. Poichè $f \in C^2(\mathrm{Int}(\Omega))$ possiamo considerare i punti critici e per ciascuno di questi, consideriamo x_0, esaminare la funzione f ristretta ad una palla, eventualmente di raggio piccolo, per cui $B(x_0, r_{x_0}) \subset \Omega$. Notiamo che $f_{|B(x_0,r_{x_0})} \in C^2(B(x_0, r_{x_0}))$. Sono inoltre soddisfatte le ipotesi in base alle quali possiamo applicare la formula di Taylor con centro x_0. Avremo

$$f(x) = f(x_0) + \frac{1}{2}\langle Hf(x_0)(x - x_0), (x - x_0)\rangle + o(||x - x_0||^2), \;\; x \to x_0,$$

ricordiamo che x_0 è critico, quindi $\nabla f(x_0) = 0$. Inoltre se $x \neq x_0$

$$f(x) - f(x_0) = (||x - x_0||^2)(\frac{1}{2}\langle Hf(x_0)(\frac{x - x_0}{||x - x_0||}), (\frac{x - x_0}{||x - x_0||})\rangle + o(1)), \;\; x \to x_0.$$

Ricordando (1) e (2) avremo

$$\frac{1}{2}\lambda \leq \frac{1}{2}\langle Hf(x_0)\frac{x - x_0}{||x - x_0||}, \frac{x - x_0}{||x - x_0||}\rangle \leq \frac{1}{2}\Lambda$$

dove λ e Λ sono rispettivamente il minimo e il massimo degli autovalori della matrice $Hf(x_0)$. Quindi se $Hf(x_0)$ è definita positiva, allora $\lambda > 0$ (e viceversa), quindi

$$f(x) - f(x_0) = (||x - x_0||^2)(\frac{1}{2}\lambda + o(1)), \;\; x \to x_0$$

e dunque per il teorema della permanenza del segno, esiste una palla centrata in x_0 e di raggio ρ tale che

$$(||x - x_0||^2)(\frac{1}{2}\lambda + o(1)) > 0$$

per ogni $x \in B(x_0, \rho) \setminus \{x_0\}$. Come conseguenza avremo che anche

$$f(x) - f(x_0) > 0$$

per ogni $x \in B(x_0, \rho) \setminus \{x_0\}$. Ovvero x_0 è punto di minimo relativo forte. Analogamente si procede nel caso in cui la matrice è definita negativa.

Remark 3.1. *Il precedente risultato è inefficace nel caso in cui x_0 è critico e $Hf(x_0) = 0$ o anche soltanto semidefinita (positiva o negativa).*