Laboratorio di Matematica, A.A. 2009-2010; I modulo;

Lezione XIII.

1. Sia $f : A \to \mathbb{R}$ una funzione definita su un insieme A contenuto in $\mathbb{R}$, e sia x^* un punto di $\mathbb{R}$. Supponiamo che A contenga un intervallo aperto centrato in x^*, privato del suo centro x^*. Sia l un punto in $\mathbb{R}$.

   Se per ogni intervallo $I(l)$ centrato in l esiste un intervallo $I^\circ(x^*)$ centrato in x^* e privato di x^* tale che
   $$f\left(I^\circ(x^*)\right) \subseteq I(l),$$
   allora diciamo che la funzione f tende a l per x che tende ad x^*, e scriviamo
   $$f(x) \to l \textit{ per } x \to x^*.$$
   In altri termini, cio' accade se per ogni $\varepsilon > 0$ esiste un $\delta > 0$ tale che per ogni $x \in A$ con $0 < |x - x^*| < \delta$ si abbia
   $$|f(x) - l| < \varepsilon.$$
   Questa disuguaglianza si puo' riscrivere come
   $$l - \varepsilon < f(x) < l + \varepsilon.$$
   Proposizione 1. *Siano f, A, x^*, l come sopra, e sia $f(x) \to l$ per $x \to x^*$. Se $l > 0$, allora esiste un $\delta > 0$ tale che che per ogni $x \in A$ con $0 < |x - x^*| < \delta$ si abbia $f(x) > 0$.*

   Questa proposizione vine detta *principio di permanenza del segno.* Essa segue direttamente dalla definizione di limite. Infatti prendendo $\varepsilon = \frac{l}{2}$, si ha che esiste un $\delta > 0$ tale che che per ogni $x \in A$ con $0 < |x - x^*| < \delta$ si abbia
   $$l - \frac{l}{2} < f(x) < l + \frac{l}{2};$$
   l'estremo sinistro e' $\frac{l}{2} > 0$, cosi' anche $f(x) > 0$.

2. **Teorema 1.** *Sia $f : I \to \mathbb{R}$ una funzione definita su un intervallo aperto A contenuto in $\mathbb{R}$, e sia $x^* \in I$; supposto che f sia di classe $\mathcal{C}^2$ su I, si ha: se $f'(x^*) = 0$ e*

   $f''(x^*) > 0$,

   allora x^* e' un punto di minimo relativo forte per f.

   Dimostrazione Per l'ipoesi f di classe $\mathcal{C}^2$ su I, si ha che f si puo' bene approssimare vicino a x^* con un polinomio di Taylor del II ordine:
   $$f(x^* + h) = f(x^*) + f'(x^*)h + \frac{1}{2}f''(x^*)h^2 + R_2(h), \tag{1}$$

con
$$\frac{R_2(h)}{h^2} \to 0 \quad per \quad h \to 0. \tag{2}$$
Ora, essendo $f'(x^*) = 0$, la (1) diviene
$$f(x^* + h) = f(x^*) + \frac{1}{2} f''(x^*)h^2 + R_2(h),$$
che puo' essere riscritta come
$$f(x^* + h) - f(x^*) = \frac{1}{2} f''(x^*)h^2 + R_2(h).$$
Per $h > 0$, possiamo dividere entrambe i membri per h^2, ottenendo cosi'
$$\frac{f(x^* + h) - f(x^*)}{h^2} = \frac{1}{2} f''(x^*) + \frac{R_2(h)}{h^2}.$$
Passando al limite per $h \to 0$ entrambe i membri, dalla (2) si ottiene
$$lim_{h\to 0} \frac{f(x^* + h) - f(x^*)}{h^2} = lim_{h\to 0} \left(\frac{1}{2} f''(x^*) + \frac{R_2(h)}{h^2} \right) = \frac{1}{2} f''(x^*).$$
Per ipotesi e' $f''(x^*) > 0$. Dunque, per il principio di permanenza del segno, si ha che esiste un $\delta > 0$ tale che che per ogni $h \in \mathbb{R}$ con $0 < |h| < \delta$ si abbia
$$\frac{f(x^* + h) - f(x^*)}{h^2} > 0,$$
cioe'
$$f(x^* + h) - f(x^*) > 0.$$
Cio' significa che x^* e' un punto di minimo forte per f in I.

3. Sia $f : A \to \mathbb{R}$ una funzione definita su un insieme A contenuto in $\mathbb{R}^n$, e sia x^* un punto di $\mathbb{R}$. Supponiamo che A contenga una sfera aperta centrata in x^*, privata del suo centro x^*. Sia l un punto in $\mathbb{R}$.

   Se per ogni intervallo $I(l)$ centrato in l esiste una sfera aperta $B^\circ(x^*)$ centrata in x^* e privata di x^* tale che
$$f\left(B^\circ(x^*)\right) \subseteq I(l),$$
   allora diciamo che la funzione f tende a l per x che tende ad x^*, e scriviamo
$$f(x) \to l \; per \; x \to x^*.$$
   In altri termini, cio' accade se per ogni $\varepsilon > 0$ esiste un $\delta > 0$ tale che per ogni $x \in A$ con $0 < \|x - x^*\| < \delta$ si abbia
$$|f(x) - l| < \varepsilon.$$
   Questa disuguaglianza si puo' riscrivere come
$$l - \varepsilon < f(x) < l + \varepsilon.$$
   Il principio di permanenza del segno val anche in questo caso, e si dimostra allo stesso modo.

Proposizione 2. *Siano f, A, x^*, l come sopra, e sia $f(x) \to l$ per $x \to x^*$. Se $l > 0$, allora esiste un $\delta > 0$ tale che che per ogni $x \in A$ con $0 < \|x - x^*\| < \delta$ si abbia $f(x) > 0$.*

4.

Teorema 2. *Sia $f : A \to \mathbb{R}$ una funzione definita su un aperto A contenuto in $\mathbb{R}$, e sia $x^* \in A$; supposto che f sia di classe $\mathcal{C}^2$ su A, si ha: se $\nabla f(x^*) = 0$ e*

$D^2 f(x^*) \quad definita\ positiva,$

allora x^* e' un punto di minimo relativo forte per f.

Dimostrazione Per l'ipoesi f di classe $\mathcal{C}^2$ su A, si ha che f si puo' bene approssimare vicino a x^* con un polinomio di Taylor del II ordine:

$$f(x^* + h) = f(x^*) + \nabla f(x^*) \cdot h + \frac{1}{2} h^T D^2 f(x^*) h + R_2(h), \tag{3}$$

con

$$\frac{R_2(h)}{\|h\|^2} \to 0 \quad per \quad h \to 0. \tag{4}$$

Ora, essendo $\nabla f(x^*) = 0$, la (3) diviene

$$f(x^* + h) = f(x^*) + \frac{1}{2} h^T D^2 f(x^*) h + R_2(h),$$

che puo' essere riscritta come

$$f(x^* + h) - f(x^*) = \frac{1}{2} h^T D^2 f(x^*) h + R_2(h).$$

Per $h > 0$, possiamo dividere entrambe i membri per $\|h\|^2$, ottenendo cosi'

$$\frac{f(x^* + h) - f(x^*)}{\|h\|^2} = \frac{1}{2} \left(\frac{h}{\|h\|} \right)^T D^2 f(x^*) \left(\frac{h}{\|h\|} \right) + \frac{R_2(h)}{\|h\|^2}. \tag{5}$$

Posto $k = \frac{h}{\|h\|}$, si ha $\|k\| = 1$. Consideriamo dunque la forma quadratica

$$Q(k) = k^T D^2 f(x^*) k$$

come una funzione $Q : C \to \mathbb{R}$, dove

$$Q = \{k \in \mathbb{R}^n : \ \|k\| = 1\}.$$

Q e' una funzione continua su $\mathbb{R}^n$, C e' un insieme chiuso e limitato, dunque per il teorema di Weierstrass Q possiede punti di minimo assoluto e punti di massimo assoluto su C. In particolare, esiste un $k_0 \in C$ tale che

$$Q(k_0) \le Q(k), \qquad \forall k \in C.$$

Per questa disuguaglianza, otteniamo dalla (5) la disuguaglianza

$$\frac{f(x^* + h) - f(x^*)}{\|h\|^2} \ge \frac{1}{2} Q(k_0) + \frac{R_2(h)}{\|h\|^2}. \tag{6}$$

Passando al limite per $h \to 0$ il secondo membro di questa disuguaglianza, per la (4) si ottiene

$$lim_{h\to 0}\left(\frac{1}{2}Q(k_0)+\frac{R_2(h)}{\|h\|^2}\right)=\frac{1}{2}Q(k_0).$$

Per ipotesi, abbiamo che $D^2f(x^*)$ e' definita positiva, cosi' $Q(k_0) > 0$. Per il principio di permanenza del segno, si ha che esiste un $\delta > 0$ tale che che per ogni $h \in \mathbb{R}^n$ con $0 < \|h\| < \delta$ si abbia

$$Q(k_0)+\frac{R_2(h)}{\|h\|^2} > 0.$$

Per questa disuguaglianza, otteniamo dalla (6), sotto la condizione $0 < \|h\| < \delta$, la disuguaglianza

$$\frac{f(x^*+h)-f(x^*)}{\|h\|^2} > 0,$$

cioe'

$$f(x^*+h)-f(x^*) > 0.$$

Cio' significa che x^* e' un punto di minimo forte per f in A.

5. Per brevita', in questo paragrafo useremo la seguente notazione per le matrici: indicheremo una matrice con una lettera minuscola, o con una parola, e indicheremo gli elementi della matrice con la stessa lettera, o parola, munita di due indici. Cosi' indicheremo una generica matrice di tipo $m \times n$, una generica matrice di tipo $n \times p$, e la matrice di tipo $m \times p$ loro prodotto con

$$a=[a_{ij}]_{\substack{i=1,\dots,m\\ j=1,\dots,n}}, \qquad b=[b_{ij}]_{\substack{j=1,\dots,n\\ h=1,\dots,p}}, \qquad ab=[(ab)_{ih}]_{\substack{i=1,\dots,m\\ h=1,\dots,p}},$$

dove

$$(ab)_{ih}=\sum_{j=1}^{n} a_{ij}b_{jh}, \qquad \forall i,h.$$

Per le trasposte avremo

$$a^T=[(a^T)_{ji}]_{\substack{j=1,\dots,n\\ i=1,\dots,m}},$$

dove

$$(a^T)_{ji}=a_{ij}, \qquad \forall i,j.$$

Sia dato un sistema

$$\sum_{j=1}^{n} a_{ij}x_j=b_i, \qquad i=1,\dots,n$$

di m equazioni lineari in n incognite $x_1,\dots,x_n$, che possiamo indicare in breve con

$$ax=b,$$

dove a e' la matrice $m \times n$ dei coefficienti, x la colonna delle n incognite, e b la colonna degli m termini noti.

Supponiamo che il numero m delle equazioni sia molto superiore al numero delle incognite, e che il sistema non abbia alcuna soluzione (questa e' una situazione molto comune in certe applicazioni).

Sia v valore di x in $\mathbb{R}^n$. Per ogni $i = 1, \ldots, m$ possiamo considerare nella $i-$ma equazione lo scarto fra il valore del primo membro su v e il secondo

$$\sum_{j=1}^{n} a_{ij} v_j - b_i,$$

e con la somma dei quadrati di questi scarti

$$\sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} v_j - b_i \right)^2$$

possiamo misurare l'errore che commettiamo assumendo che v sia una soluzione. Diremo che un vettore v^* di $\mathbb{R}^n$ e' una *soluzione ai minimi quadrati* del sistema se l'errore corrispondente a v^* e' minimo fra tutti gli errori corrispondenti ai vettori v di $\mathbb{R}^n$.

Associando a ogni vettore di $\mathbb{R}^n$ il corrispondente errore abbiamo una funzione $E : \mathbb{R}^n \to \mathbb{R}$, definita appunto dal polinomio

$$E(x_1, \ldots, x_n) = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} x_j - b_i \right)^2,$$

di secondo grado in $x_1, \ldots, x_n$. Le soluzioni ai minimi quadrati del sistema sono i punti di minimo assoluto della funzione E.

Le derivate parziali di E sono date da

$$D_{x_h} \left(\sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} x_j - b_i \right)^2 \right) = \sum_{i=1}^{m} D_{x_h} \left(\sum_{j=1}^{n} a_{ij} x_j - b_i \right)^2$$

$$\sum_{i=1}^{m} 2a_{ih} \left(\sum_{j=1}^{n} a_{ij} x_j - b_i \right) = 2 \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ih} a_{ij} x_j - 2 \sum_{i=1}^{m} a_{ih} b_i.$$

I punti di minimo assoluto devono essere punti stazionari per E :

$$D_{x_h} E(x) = 0, \qquad h = 1, \ldots, n,$$

cioe' soluzioni del sistema

$$\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ih} a_{ij} x_j = \sum_{i=1}^{m} a_{ih} b_i, \qquad h = 1, \ldots, n,$$

di n equazioni lineari nelle n incognite $x_1, \ldots, x_n$.

Questo sistema si puo' scivere nella forma

$$\sum_{i=1}^{m} \sum_{j=1}^{n} (a^T)_{hi} a_{ij} x_j = \sum_{i=1}^{m} (a^T)_{hi} b_i, \qquad h = 1, \dots, n,$$

poi come

$$(a^T ax)_h = (a^T b)_h, \qquad h = 1, \dots, n,$$

cioe'

$$a^T ax = h.$$

Supponiamo che la matrice a ha rango pieno; allora la matrice $a^T a$ e' definita positiva e dunque invertibile.

C'e' uno ed un solo punto stazionario dato da

$$x^* = (a^T a)^{-1} b.$$

La matrice Hessiana di E e' data da

$$D^2 E(x) = a^T a,$$

che e' definita positiva, per ogni x in $\mathbb{R}^n$.

Dunque x^* e' un punto di minimo relativo forte per E. Si puo' provare, usando ad esempio le prime proprieta' delle funzioni convesse, che e' anche un punto di minimo assoluto per E.

Riferimenti.

I punti 1,2,3 sono preparatori al punto 4.

per il punto 4: SB, Cap. 7, par. 7.6, p.174-175.

Il punto 5 non compare sul testo; per una sua applicazione: SB, Cap. 7, par. 7.5, p.170-172