LA FORMULA DI TAYLOR PER FUNZIONI DI PIÙ VARIABILI

L'arsenale degli artifici per estendere a più variabili risultati stabiliti per funzioni d'una sola variabile reale riserva un posto d'onore a quello, semplicissimo, di considerare la restrizione della funzione f da studiare ad un segmento contenuto nel dominio $\gamma(t) = x_0 + tv$, il che ne trasforma lo studio in quello di $h(t) = f(\gamma(t))$, funzione d'una sola variabile t. La condizione necessaria per un estremo (di Fermat), stabilita direttamente in una variabile, viene così estesa alle funzioni tanto su $\mathbb{R}^n$, quanto negli spazi di dimensione infinita (Equazione di Eulero, o di Eulero-Lagrange per i meccanici razionali).

Lo stesso "trucco" svolge un ottimo lavoro per generalizzare a più variabili il naturale risultato di Taylor, che verrà qui esposto, per semplicità, col resto nella forma "di Lagrange". Ricordiamone brevemente l'enunciato.

Sia $f \in C^{N+1}[x_0, x]$. Allora, esiste $\xi \in \,]x_0, x[$ ~~tale~~ che

$$f(x) = \sum_{k=0}^{N} \frac{1}{k!} f^{(k)}(x_0)(x-x_0)^k + \frac{f^{(N+1)}(\xi)}{(N+1)!}(x-x_0)^{N+1}$$

Si ricordi che esistono enunciati più raffinati con ipotesi meno restrittive (reperibili, ad esempio, su G. PRODI Analisi Matematica ed. BORINGHIERI). E' bene ricordare anche che $f \in C^{N+1}[a,b]$ vuol dire che f ha tutte le derivate continue fino all'ordine $N+1$ non solo nei punti interni $]a,b[$, ma anche in a e in b! Naturalmente, le derivate in a saranno destre e quelle in b sinistre. Un modo rapido di verificare tali ipotesi è di stabilire che f è $C^{N+1}]a',b'[$ con $]a',b'[\supset [a,b]$. Inoltre, ricordiamo che $0! = 1$, che $f^{(0)}(x_0) = f(x_0)$, e che la sommatoria è un polinomio, detto di Taylor di grado N, di f in x_0.

Nel caso $N=0$, il teorema di Taylor produce un risultato altrimenti ben noto: $f(x) = f(x_0) + f'(\xi)(x-x_0)$, il che giustifica la denominazione "resto di Lagrange" per l'ultimo termine della formula nell'enunciato precedente.

Verranno presentate due diverse espressioni del polinomio di Taylor in più variabili: la prima deriva direttamente da quella in una variabile, mediante l'artificio descritto più su, mentre l'altra comporta un numero minore di calcoli di derivate, perché tiene conto del teorema di Schwartz.

Sia dunque $f: \Omega \to \mathbb{R}$, $\Omega \subseteq \mathbb{R}^n$.

<u>TEOREMA</u>: <u>Sia</u> $f \in C^{N+1}(B_\delta(x_0))$. <u>Allora</u>, <u>per ogni</u> w verificante $|w|<\delta$ <u>esiste</u> $\xi \in]0,1[$ <u>tale che</u>

$$f(x_0+w) = \sum_{k=0}^{N} \frac{1}{k!} \sum_{i_1,\dots,i_k=1}^{n} f_{x_{i_1} x_{i_2} \cdots x_{i_k}}(x_0)\, w_{i_1} w_{i_2} \cdots w_{i_k} +$$

$$+ \frac{1}{(N+1)!} \sum_{i_1,\dots,i_{N+1}=1}^{n} f_{x_{i_1} \cdots x_{i_{N+1}}}(x_0+\xi w)\, w_{i_1} \cdots w_{i_{N+1}}$$

Il primo addendo a secondo membro, con le due sommatorie annidate, è il polinomio di Taylor. Nell'enunciato precedente si è posto $w = x - x_0$ per ottenere maggiore semplicità nell'espressione a secondo membro.

<u>DIM</u>. Fissato ad arbitrio $x \in B_\delta(x_0)$, e posto $w = x - x_0$ si ha $|w|<\delta$. Si pone, allora, $h(t) = f(x_0+tw)$ e si osserva che $h(0) = f(x_0)$ e $h(1) = f(x)$.

Verrà tra breve verificato che $h \in C^{N+1}[0,1]$ sicché, dal teorema in una variabile, ne segue che $\exists\, \xi \in]0,1[$ tale che

$$h(1) = h(0) + h'(0)\cdot 1 + \dots + \frac{1}{N!} h^{(N)}(0)\, 1^N + \frac{1}{(N+1)!} h^{(N+1)}(\xi)\, 1^{N+1} =$$

$$= \sum_{k=0}^{N} \frac{1}{k!} h^{(k)}(0) + \frac{1}{(N+1)!} h^{(N+1)}(\xi)$$

Occorre dunque calcolare le derivate successive $h^{(0)}(t), h^{(1)}(t), \dots$
$\dots, h^{(N+1)}(t)$. Osserviamo che $h(t)$ è definita su $[-1, 1]$.
Si ha subito $h^{(0)}(t) = h(t) = f(x_0 + tw)\big|_{t=0} = f(x_0)$
Per calcolare $h^{(1)}(t) = \frac{d}{dt}\left[f(x_0+tw)\right]$ si può utilizzare il teorema di derivazione delle funzioni composte, perché f è (molto più che) C^1 in un intorno di 0 ed è quindi differenziabile, così come lo è $t \to x_0 + tw$. Ne segue

$$h'(t) = \frac{d}{dt}\left[f(x_0+tw)\right] = \sum_{i_1=1}^{n} f_{x_{i_1}}(x_0+tw)\, w_{i_1}$$

da cui

$$h'(0) = \sum_{i_1=1}^{n} f_{x_{i_1}}(x_0)\, w_{i_1}$$

Si può calcolare $h''(t)$ riapplicando il teorema di derivazione delle funzioni composte a ciascuno dei termini $f_{x_{i_1}}(x_0+tw)\, w_{i_1}$ e, perché le derivate considerate hanno a loro volta derivate continue, esse saranno differenziabili e il teorema potrà essere applicato. Ne segue che

$$h''(t) = \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} f_{x_{i_1} x_{i_2}}(x_0+tw)\, w_{i_1} w_{i_2}$$

$$h'''(t) = \sum_{i_1, i_2, i_3=1}^{n} f_{x_{i_1} x_{i_2} x_{i_3}}(x_0+tw)\, w_{i_1} w_{i_2} w_{i_3}$$

$$\vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad \vdots$$

$$h^{(N)}(t) = \sum_{i_1, i_2, \dots, i_N = 1}^{n} f_{x_{i_1} x_{i_2} \dots x_{i_N}}(x_0 + tw)\, w_{i_1} \dots w_{i_N}$$

Ponendovi $t=0$, si ottengono i coefficienti della formula di Taylor per h

$$h^{(0)}(0) = f(x_0)$$

$$h'(0) = \sum_{i_1=1}^{n} f_{x_{i_1}}(x_0)\, w_{i_1}$$

$$h''(0) = \sum_{i_1, i_2 = 1}^{n} f_{x_{i_1} x_{i_1}}(x_0)\, w_{i_1} w_{i_2}$$

$$\vdots \qquad\qquad \vdots$$

$$h^{(N)}(0) = \sum_{i_1, \dots, i_N = 1}^{n} f_{x_{i_1} \dots x_{i_N}}(x_0)\, w_{i_1} \dots w_{i_N}$$

da cui, infine, per un opportuno $\xi \in \,]0,1[$

$$f(x) = h(1) = \sum_{k=1}^{N} \frac{1}{k!} h^{(k)}(0) + \frac{1}{(N+1)!} h^{(N+1)}(\xi) =$$

$$= \sum_{k=0}^{N} \frac{1}{k!} \sum_{i_1 \dots i_k = 1}^{n} f_{x_{i_1} \dots x_{i_k}}(x_0)\, w_{i_1} \dots w_{i_k} +$$

$$+ \frac{1}{(N+1)!} \sum_{i_1, \dots, i_{N+1} = 1}^{n} f_{x_{i_1} \dots x_{i_{N+1}}}(x_0 + \xi w)\, w_{i_1} \dots w_{i_{N+1}}$$

E' utile osservare che il teorema appena provato, seppure con ipotesi leggermente più restrittive, ($f \in C^{N+1}$), fornisce anche una stima del resto di tipo "Peano":

Infatti, il modulo del resto

$$\frac{1}{(N+1)!}\left|\sum_{i_1 \dots i_{N+1}=1}^{n} f_{x_{i_1} \dots x_{i_{N+1}}}(x_0+\xi w)\, w_{i_1} \dots w_{i_{N+1}}\right| \equiv |R_N(w)|$$

può essere stimato per la disuguaglianza triangolare, con

$$|R_N(w)| \le \sum_{i_1 \dots i_{N+1}=1}^{n} \max_{x \in B_{\delta'}(x_0)} \left|f_{x_{i_1} \dots x_{i_{N+1}}}(x)\right| \left|w_{i_1} \dots w_{i_{N+1}}\right| \equiv S_N(w)$$

da cui, essendo $S_N(w)$ una funzione $(N+1)$-omogenea, ne segue subito che

$$\lim_{w \to 0} \frac{S_N(w)}{|w|^N} = 0 \qquad (\underline{\text{PEANO}})$$

in quanto $S_N(w)/|w|^N$ è 1-omogenea e limitata sulla sfera unitaria $|w|=1$, essendo ivi continua.

Utilizzando tale stima "puntuale" sull'ordine di infinitesimo del resto (tipica del resto di Peano), si possono estendere a più variabili le condizioni sufficienti per un estremo locale per le funzioni di una variabile, basate sulle derivate seconde. Verrà provato il:

TEOREMA Sia $f \in C^3(B_\delta(x_0))$ e sia x_0 tale che $\nabla f(x_0) = 0$. Allora, il segno di $f(x) - f(x_0)$ coincide col segno della forma quadratica, detta Hessiana, $H(x-x_0) = \sum_{i,j=1}^{n} f_{x_i x_j}(x_0)(x-x_0)_i (x-x_0)_j$, purché $|x - x_0|$ sia sufficientemente piccolo.

E' bene notare che la condizione $\nabla f(x_0) = 0$ è quella necessaria (di Fermat) per i massimi e i minimi, mentre le condizioni sufficienti si possono ricavare dallo studio del segno della forma quadratica H, e sono legate al segno dei suoi autovalori: più complicato che in una variabile, ma non di molto!

DIM. Sia $x \in B_\delta(x_0)$ e si ponga, per semplicità, $w = x - x_0$.

Poiché $f \in C^3$ si può scrivere la formula di Taylor arrestandosi al grado $N = 2$, con il resto di tipo "Peano", infinitesimo d'ordine superiore rispetto a $|w|^2$.

$$f(x_0+w) = f(x_0) + \nabla f(x_0) w + \frac{1}{2!}\sum_{i,j=1}^{n} f_{x_i x_j}(x_0) w_i w_j + R_2(w)$$

ove $\lim_{w\to 0} R_2(w)/|w|^2 = 0$

Poiché $\nabla f(x_0)=0$, si ha

$$f(x_0+w)-f(x_0) = |w|^2\left[\frac{1}{2}\,\frac{\sum_{i,j=1}^{n} f_{x_i x_j}(x_0) w_i w_j}{|w|^2} + \frac{R_2(w)}{|w|^2}\right] \quad (*)$$

Dunque, il segno di $f(x_0+w)-f(x_0)$ è correlato a quello dell'espressione in parentesi quadra ($|w|^2 \geq 0$).
Osserviamo ora che il primo addendo è una funzione g 0-omogenea (rapporto di funzioni 2-omogenee), ovunque definita sulla sfera unitaria $|w|=1$, ed ha quindi su di essa minimo λ e massimo Λ (globali). Poiché è 0-omogenea, tali valori sono estremi anche in tutto $\mathbb{R}^n\setminus\{0\}$, perché per ogni $u\neq 0$ si ha

$$g(u) = g\left(|u|\frac{u}{|u|}\right) = |u|^0 g\left(\frac{u}{|u|}\right) = g\left(\frac{u}{|u|}\right)$$

e dunque

$$\lambda \leq g(u) \leq \Lambda \qquad \forall u\in\mathbb{R}^n\setminus\{0\}$$

e poiché tali valori sono assunti su $\{|u|=1\}\subseteq\mathbb{R}^n\setminus\{0\}$ essi sono estremi globali in tutto $\mathbb{R}^n-\{0\}$.

L'algebra lineare offre un quadro completo del segno di $\sum f_{x_i x_j}(x_0) w_i w_j$, che risulta

1) strettamente positiva in $\mathbb{R}^n \setminus \{0\}$ se e solo se tutti gli autovalori sono (tutti) strettamente positivi.

2) strettamente negativa se e solo se tutti gli autovalori sono strettamente negativi.

Caso 1) La forma $\sum f_{x_i x_j}(x_0) w_i w_j$, i cui coefficienti formano la **matrice Hessiana di f**, ha tutti gli autovalori strettamente positivi, e quindi è sempre strettamente positiva sulla sfera unitaria, da cui $\boxed{\lambda > 0}$.

Allora, scelto $\varepsilon = \frac{\lambda}{2}$ $\exists \rho$: se $w \in B_\rho(0)$ si ha

$$\left| \frac{R_2(w)}{|w|^2} \right| < \frac{\lambda}{2}$$

da cui se $0 < |w| < \rho$

$$\frac{1}{2} \underbrace{\frac{\sum f_{x_i x_j}(x_0) w_i w_j}{|w|^2}}_{\geq \lambda > 0} + \overbrace{\frac{R_2(w)}{|w|^2}}^{\in \left]-\frac{\lambda}{2}, \frac{\lambda}{2}\right[} > 0$$

e, di conseguenza, $f(x_0 + w) - f(x_0) > 0$ se $|w| < \rho$, e dunque x_0 è di minimo locale.

Analogamente si ragiona nel caso 2), quando tutti gli autovalori della matrice Hessiana siano negativi.

2) Stavolta $\Lambda < 0$, e scelto $\varepsilon = \frac{|\Lambda|}{2}$

si otterrà un intorno $B_\rho(0)$ nel quale $\left| \frac{R_2(w)}{|w|^2} \right| < \frac{|\Lambda|}{2}$, da cui

$$\underbrace{\frac{1}{2} \frac{\sum f_{x_i x_j}(x_0) w_i w_j}{|w|_2}}_{\leq \Lambda < 0} + \underbrace{\frac{R_2(w)}{|w|^2}}_{\in \left]-\frac{|\Lambda|}{2}, \frac{|\Lambda|}{2}\right[} < 0$$

e, infine, $f(x_0 + w) - f(x_0) \leq 0$ in $B_\rho(0)$.

Ne seguono dunque le due condizioni sufficienti.

> Un punto critico x_0 $(\nabla f(x_0) = 0)$, in cui la forma Hessiana è definita positiva è un minimo locale se è definita negativa è un massimo locale.

La forma Hessiana può presentare, come è noto dall'Algebra, altri tre casi:

- un autovalore nullo e tutti gli altri positivi
- un autovalore nullo e tutti gli altri negativi
- due autovalori discordi (non nulli).

I primi due casi sono estremamente ostici: infatti, è vero che il complesso dei termini di secondo grado, quando non si annulla, ha segno costante; il problema è che nelle direzioni nelle quali si annulla il segno dipenderà dal resto, che non è più trascurabile. Tutto diventa chiarissimo con un esempio (anzi: due!)

$$f(x,y) = x^2 + y^4 \qquad g(x,y) = x^2 - y^4$$

Si vede subito che $(0,0)$ è di minimo (globale) per f mentre non lo è per g, che è negativa sull'asse y e positiva sull'asse x, e 0 in $(0,0)$. Ciononostante,

$$H(f) = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 12y^2 \end{pmatrix} \Bigg|_{(0,0)} = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$$

diagonali

autovalori $\underline{2 \text{ e } 0}$

mentre

$$H(g) = \begin{pmatrix} g_{xx} & g_{xy} \\ g_{yx} & g_{yy} \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & -12y^2 \end{pmatrix} \Bigg|_{(0,0)} = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$$

e dunque se l'Hessiana è semi-definita, NULLA può dirsi senza ulteriori indagini sui termini di grado superiore del polinomio di Taylor.

Rimane ancora fuori il caso dell'Hessiana indefinita.

In tal caso $\sum f_{x_i x_j}(x_0) w_i w_j$ cambia segno e dunque esistono u, v, con $|u| = |v| = 1$, tali che $g(u) > 0$ e $g(v) < 0$.

Ragionando come prima, scegliendo rispettivamente $\varepsilon = g(u)/2$ o $\varepsilon = |g(v)|/2$ si ottiene $f(x_0 + tu) - f(x_0) > 0$ e $f(x_0 + sv) - f(x_0) < 0$ per ogni t, s di modulo abbastanza piccolo. Ne segue che x_0 non può essere né di massimo né di minimo locale perché in ogni sfera comunque piccola cadranno punti del tipo $x_0 + tu$, ai quali f assume valori maggiori di $f(x_0)$, ma anche punti del tipo $x_0 + sv$, ai quali f assume valori minori.

Un punto x_0 critico ($\nabla f(x_0) = 0$) con Hessiana indefinita non è né di massimo né di minimo ed è detto DI SELLA (non degenere).

<u>RIASSUMENDO</u>: Se $\nabla f(x_0) = 0$, allora:

1) <u>Hessiana definita positiva</u> $\Rightarrow x_0$ <u>minimo locale</u>
2) <u>Hessiana definita negativa</u> $\Rightarrow x_0$ <u>massimo locale</u>
3) <u>Hessiana indefinita</u> $\Rightarrow x_0$ <u>di sella</u> (né max, né min)

4) <u>Hessiana semidefinita</u> $\Rightarrow$ <u>NULLA PUÒ DIRSI</u>

Si noti subito che il caso 4), l'Hessiana semidefinita, che corrisponde in più variabili al caso critico $f''(x_0)=0$ in una variabile, e per la stessa ragione. Infatti,

$$f(x)-f(x_0)=\frac{1}{2}f''(x_0)(x-x_0)^2+R_2(x-x_0)$$

e, se $f''(x_0)=0$, il segno di $f(x)-f(x_0)$ è determinato unicamente dal segno del resto $R_2(x-x_0)$, imprevedibile senza ulteriori indagini sui termini di ordine superiore.

Ciò è esattamente quello che succede nelle direzioni w del nucleo dell'Hessiana, lungo le quali il complesso dei termini di secondo grado si annulla, lasciando il campo a quelli d'ordine superiore, così come accade in una variabile.

<u>NOTA</u>. Se (a_{ij}) è una matrice simmetrica, si può provare che gli estremi di $\sum a_{ij}w_iw_j/|w|^2$ sono esattamente il massimo ed il minimo autovalore Λ e λ di (a_{ij}) e, dunque

$$\lambda|w|^2\leq\sum a_{ij}w_iw_j\leq\Lambda|w|^2\quad\forall w\in\mathbb{R}^n$$

Grazie al teorema di Schwartz, l'Hessiana $\left(f_{x_ix_j}(x_0)\right)$ è di certo simmetrica nelle ipotesi prima introdotte $(f\in C^3)$.