Formula di Taylor e multi-indici

Dato un aperto $\Omega \subset \mathbb{R}^N$ e supposta $f : \Omega \to \mathbb{R}$ (almeno) differenziabile n volte in un punto $\boldsymbol{x}_0 \in \Omega$, ricordo la formula di Taylor di ordine m:

$$f(\boldsymbol{x}) = T_m(\boldsymbol{x}) + o(|\boldsymbol{x} - \boldsymbol{x}_0|^m) \quad \text{per } \boldsymbol{x} \to \boldsymbol{x}_0, \tag{1}$$

dove T_m è il polinomio di Taylor di ordine m della funzione f con centro $\boldsymbol{x}_0$. Quest'ultimo si può esprimere decomponendolo come somma di termini omogenei di grado n, $n = 0, \ldots, m$, nel seguente modo

$$T_m(\boldsymbol{x}) = \sum_{n=0}^{m} S_n(\boldsymbol{x}).$$

Posto per comodità $\boldsymbol{h} := \boldsymbol{x} - \boldsymbol{x}_0$, osservo per esempio che

$$S_0(\boldsymbol{x}) = f(\boldsymbol{x}_0), \quad S_1(\boldsymbol{x}) = (\nabla f)(\boldsymbol{x}_0) \cdot \boldsymbol{h}, \quad S_2(\boldsymbol{x}) = \frac{1}{2}\langle Hf(\boldsymbol{x}_0)\boldsymbol{h}, \boldsymbol{h}\rangle. \tag{2}$$

Il termine omogeneo di ordine $n > 2$ è più difficile da scrivere per esteso. Infatti è necessario sommare rispetto a n indici distinti $i_1, \ldots, i_n$, ciascuno dei quali varia da 1 a N, ove N è la dimensione di Ω. Si ha, cioè,

$$S_n(\boldsymbol{x}) = \frac{1}{n!} \sum_{i_1,\ldots,i_n=1}^{N} \frac{\partial^n f(\boldsymbol{x}_0)}{\partial x_{i_1} \ldots \partial x_{i_n}} h_{i_1} \ldots h_{i_n}. \tag{3}$$

Per verificare l'esattezza di questa scrittura si può confrontare con l'espressione esplicita del termine in (2) dipendente dalla matrice Hessiana (ovvero S_2) che, se esplicitato come somma, diviene

$$S_2(\boldsymbol{x}) = \frac{1}{2}\langle Hf(\boldsymbol{x}_0)\boldsymbol{h}, \boldsymbol{h}\rangle = \frac{1}{2!} \sum_{i,j=1}^{2} \frac{\partial^2 f(\boldsymbol{x}_0)}{\partial x_i \, \partial x_j} h_i h_j,$$

dove naturalmente i puntini sono scomparsi in quanto ho solo due ordini di derivazione anziché n.

Per esprimere in maniera più compatta la formula (3), si può ricorrere al concetto di *multi-indice*: per definizione, un multi-indice di *lunghezza* n in N variabili è un vettore $\alpha = (\alpha_1, \ldots, \alpha_N) \in \mathbb{N}^N$ tale che $|\alpha| := \alpha_1 + \cdots + \alpha_N = n$. Se, per esempio, $N = 3$ e $n = 7$, il multiindice $\alpha = (2, 1, 4)$ (la cui lunghezza è, appunto, $n = 7$ in quanto $2 + 1 + 4 = 7$) può essere utilizzato per rappresentare una derivata parziale di ordine 7, ottenuta derivando 2 volte rispetto alla prima variabile x, 1 volta rispetto alla seconda variabile y e 4 volte rispetto alla terza variabile z.

A livello di notazione generale, pongo cioè

$$D^\alpha f(\boldsymbol{x}_0) := \frac{\partial^n f(\boldsymbol{x}_0)}{\partial x_1^{\alpha_1} \ldots \partial x_N^{\alpha_N}}, \quad \text{ove, ricordo, } n = |\alpha| = \alpha_1 + \cdots + \alpha_N.$$

Per esempio, nel caso precedente $\alpha = (2, 1, 4)$, dove $N = 3$ e $n = 7$, si ha

$$D^\alpha f(\boldsymbol{x}_0) = \frac{\partial^7 f(\boldsymbol{x}_0)}{\partial x^2 \, \partial y \, \partial z^4}.$$

Introducendo inoltre (per $\boldsymbol{h} \in \mathbb{R}^N$ e $\alpha \in \mathbb{N}^N$) le notazioni

$$\alpha! := \alpha_1! \ldots \alpha_N!, \qquad \boldsymbol{h}^\alpha := h_1^{\alpha_1} \ldots h_N^{\alpha_N},$$

voglio ora utilizzare il linguaggio dei multi-indici per esprimere in maniera più concisa l'espressione (3). Inizio dunque dal caso più semplice in cui $N = 2$. In questo caso posso indicare semplicemente con x e y le componenti scalari della variabile indipendente e con $(x - x_0)$, $(y - y_0)$ le componenti scalari dell'incremento $\boldsymbol{h}$. Osservando il secondo membro di (3), mi accorgo inoltre che il simbolo di sommatoria corrisponde, nella sostanza, a una somma di 2^n addendi (dove, ricordo, n è l'ordine di derivazione); infatti ciascuno degli indici $i_1, \ldots, i_n$ può variare solamente da 1 a 2. Inoltre ciascuno di tali addendi dovrà necessariamente essere della forma

$$\frac{\partial^n f(\boldsymbol{x}_0)}{\partial^k x \, \partial^{n-k} y}(x - x_0)^k (y - y_0)^{n-k} \tag{4}$$

per qualche $k \in \mathbb{N}$, $0 \leq k \leq n$, a significare che sto derivando esattamente k volte rispetto a x e le restanti $n-k$ volte rispetto a y.

Utilizzando le nozioni di base del calcolo combinatorio, è facile vedere che, nella sommatoria a secondo membro di (3), per ogni $k = 0, \dots, n$, il termine della forma (4) si ripete esattamente $\binom{n}{k} = \dfrac{n!}{k!(n-k)!}$ volte. Per esempio, per $k = 0$, tale termine corrisponde alla derivata pura $\partial^n f/\partial y^n$, la quale compare solo una volta, ovvero quando ciascuno degli indici $i_1, \dots, i_n$ vale 2; per $k = 1$, ho invece la derivata $\partial^n f/(\partial x \partial y^{n-1})$, la quale compare $n = \binom{n}{1}$ volte (ovvero quando uno e uno solo degli indici $i_1, \dots, i_n$ assume il valore 1), e così via. Segue, dunque, che per $N = 2$, raggruppando gli addendi simili, (3) si può riscrivere nella forma

$$\begin{aligned} S_n(\boldsymbol{x}) &= \frac{1}{n!}\sum_{k=0}^{n} \frac{n!}{k!(n-k)!}\frac{\partial^n f(\boldsymbol{x}_0)}{\partial x^k\, \partial y^{n-k}}(x-x_0)^k(y-y_0)^{n-k} \\ &= \sum_{k=0}^{n} \frac{1}{k!(n-k)!}\frac{\partial^n f(\boldsymbol{x}_0)}{\partial x^k\, \partial y^{n-k}}(x-x_0)^k(y-y_0)^{n-k} \\ &= \sum_{|\alpha|=n} \frac{1}{\alpha!}D^\alpha f(\boldsymbol{x}_0)\boldsymbol{h}^\alpha, \quad \text{dove } \boldsymbol{h} = (x-x_0, y-y_0). \end{aligned} \tag{5}$$

Per dedurre l'ultima uguaglianza ho usato il fatto che ogni multiindice di lunghezza n in $\mathbb{N}^2$ ha la forma $\alpha = (k, n-k)$ per qualche $k \in \{0, \dots, n\}$ e, dunque, $\alpha! = k!(n-k)!$.

Grazie alla (5), posso inoltre notare che, almeno per $N = 2$, la formula (1) assume l'espressione "compatta"

$$f(\boldsymbol{x}) = \sum_{|\alpha| \leq m} \frac{1}{\alpha!}D^\alpha f(\boldsymbol{x}_0)(\boldsymbol{x}-\boldsymbol{x}_0)^\alpha + o(|\boldsymbol{x}-\boldsymbol{x}_0|^m) \quad \text{per } \boldsymbol{x} \to \boldsymbol{x}_0. \tag{6}$$

Mi occupo, infine, all'estensione del discorso nel caso in cui la dimensione N di Ω sia strettamente maggiore di 2, con l'obiettivo di arrivare, anche in questo caso, alla formula (6).

Per semplicità di notazione mi limito al caso $N = 3$, osservando che per $N > 3$ il discorso si può ulteriormente generalizzare procedendo per induzione su N. Ora, se $N = 3$, osservo innanzitutto che un generico multi-indice α di lunghezza n si può esprimere come $\alpha = (k, \ell, p)$ ove $k + \ell + p = n$. La sommatoria a secondo membro di (3) consiste inoltre esattamente di 3^n addendi. Tra questi, voglio vedere quanti corrispondono al dato multi-indice $\alpha = (k, \ell, p)$: per fare questo procedo attraverso un'analogia che sfrutta un classico, e forse più intuitiva, problema di calcolo combinatorio. Suppongo cioè di estrarre $n = 7$ (lunghezza del multi-indice, ovvero ordine di derivazione) volte una pallina che, in ciascuna estrazione, può essere, con uguale probabilità, verde (derivata rispetto a x), gialla (derivata rispetto a y) oppure rossa (derivata rispetto a z). In tutto ho 3^7 combinazioni (ovvero 3^7 addendi in (3)), che intendo raggruppare secondo i multi-indici α. Per esempio, se $\alpha = (7, 0, 0)$, vuol dire che ho estratto ogni volta una pallina verde (ovvero ho derivato sempre rispetto a x), e questo succede solo in un caso ($= \frac{7!}{7!0!0!}$) su 3^7. Se, invece, $\alpha = (2, 2, 3)$, vuol dire che ho estratto 2 palline verdi, 2 gialle e 3 rosse. Per vedere quante volte questo accade, osservo dapprima che le combinazioni corrispondenti a 2 palline verdi e 5 *gialle oppure rosse* sono esattamente $\binom{7}{2} = \frac{7!}{2!5!}$. Tuttavia, questa rappresentazione è incompleta: per ciascuna di queste combinazioni devo vedere in quanti sottocasi tra le 5 palline "gialle oppure rosse" ce ne sono esattamente 2 gialle e 3 rosse. Ripetendo il ragionamento mi accorgo che questo avviene esattamente $\binom{5}{2} = \frac{5!}{2!3!}$ volte. Il numero di volte in cui ho 2 palline verdi, 2 gialle e 3 rosse è dunque dato dal prodotto $\binom{7}{2}\binom{5}{2} = \frac{7!}{2!5!}\frac{5!}{2!3!} = \frac{7!}{2!2!3!}$. Generalizzando, e tornando alla formula di Taylor, nel caso di un dato multi-indice $\alpha = (k, \ell, p)$ di lunghezza n, il numero di volte in cui compare la derivata rispetto ad α nella sommatoria in (5) è dunque esattamente pari a

$$\binom{n}{k}\binom{n-k}{\ell} = \frac{n!}{k!(n-k)!}\frac{(n-k)!}{\ell!(n-k-\ell)!} = \frac{n!}{k!\ell!p!} = \frac{n!}{\alpha!},$$

che è quanto permette di arrivare di nuovo alla formula (6).