Higher-Order Derivatives and Taylor's Formula in Several Variables

G. B. Folland

Traditional notations for partial derivatives become rather cumbersome for derivatives of order higher than two, and they make it rather difficult to write Taylor's theorem in an intelligible fashion. (In particular, Apostol's $D_{r_1,\ldots,r_k}$ is pretty ghastly.) However, a better notation, which is now in common usage in the literature of partial differential equations, is available.

A **multi-index** is an n-tuple of nonnegative integers. Multi-indices are generally denoted by the Greek letters α or β:

$$\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n), \quad \beta = (\beta_1, \beta_2, \ldots, \beta_n) \qquad \big(\alpha_j, \beta_j \in \{0, 1, 2, \ldots\}\big).$$

If α is a multi-index, we define

$$|\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_n, \qquad \alpha! = \alpha_1!\alpha_2! \cdots \alpha_n!,$$
$$\mathbf{x}^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n} \text{ (where } \mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n),$$
$$\partial^\alpha f = \partial_1^{\alpha_1} \partial_2^{\alpha_2} \cdots \partial_n^{\alpha_n} f = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \cdots \partial x_n^{\alpha_n}}$$

The number $|\alpha| = \alpha_1 + \cdots + \alpha_n$ is called the **order** or **degree** of α. Thus, the order of α is the same as the order of $\mathbf{x}^\alpha$ as a monomial or the order of ∂^α as a partial derivative.

If f is a function of class C^k, by Theorem 12.13 and the discussion following it the order of differentiation in a kth-order partial derivative of f is immaterial. Thus, the generic kth-order partial derivative of f can be written simply as $\partial^\alpha f$ with $|\alpha| = k$.

Example. With $n = 3$ and $\mathbf{x} = (x, y, z)$, we have

$$\partial^{(0,3,0)} f = \frac{\partial^3 f}{\partial y^3}, \qquad \partial^{(1,0,1)} f = \frac{\partial^2 f}{\partial x \partial z}, \qquad \mathbf{x}^{(2,1,5)} = x^2 y z^5.$$

As the notation $\mathbf{x}^\alpha$ indicates, multi-indices are handy for writing not only derivatives but also polynomials in several variables. To illustrate their use, we present a generalization of the binomial theorem.

Theorem 1 (The Multinomial Theorem). *For any $\mathbf{x} = (x_1, x_2, \ldots x_n) \in \mathbb{R}^n$ and any positive integer k,*

$$(x_1 + x_2 + \cdots + x_n)^k = \sum_{|\alpha|=k} \frac{k!}{\alpha!} \mathbf{x}^\alpha.$$

Proof. The case $n = 2$ is just the binomial theorem:

$$(x_1 + x_2)^k = \sum_{j=0}^{k} \frac{k!}{j!(k-j)!} x_1^j x_2^{k-j} = \sum_{\alpha_1+\alpha_2=k} \frac{k!}{\alpha_1!\alpha_2!} x_1^{\alpha_1} x_2^{\alpha_2} = \sum_{|\alpha|=k} \frac{k!}{\alpha!} \mathbf{x}^\alpha,$$

where we have set $\alpha_1 = j$, $\alpha_2 = k - j$, and $\alpha = (\alpha_1, \alpha_2)$. The general case follows by induction on n. Suppose the result is true for $n < N$ and $\mathbf{x} = (x_1, \ldots, x_N)$. By using the result for $n = 2$ and then the result for $n = N - 1$, we obtain

$$\begin{aligned}
(x_1 + \cdots + x_N)^k &= \big[(x_1 + \cdots + x_{N-1}) + x_N\big]^k \\
&= \sum_{i+j=k} \frac{k!}{i!j!}(x_1 + \cdots + x_{N-1})^i x_N^j \\
&= \sum_{i+j=k} \frac{k!}{i!j!} \sum_{|\beta|=i} \frac{i!}{\beta!} \widetilde{\mathbf{x}}^\beta x_N^j,
\end{aligned}$$

where $\beta = (\beta_1, \ldots, \beta_{N-1})$ and $\widetilde{\mathbf{x}} = (x_1, \ldots, x_{N-1})$. To conclude, we set $\alpha = (\beta_1, \ldots, \beta_{N-1}, j)$, so that $\beta! j! = \alpha!$ and $\widetilde{\mathbf{x}}^\beta x_N^j = \mathbf{x}^\alpha$. Observing that α runs over all multi-indices of order k when β runs over all multi-indices of order $i = k - j$ and j runs from 0 to k, we obtain $\sum_{|\alpha|=k} k!\mathbf{x}^\alpha/\alpha!$. ∎

A similar argument leads to the product rule for higher-order partial derivatives:

$$\partial^\alpha (fg) = \sum_{\beta+\gamma=\alpha} \frac{\alpha!}{\beta!\gamma!} (\partial^\beta f)(\partial^\gamma g).$$

The proof is by induction on the number n of variables, the base case $n = 1$ being the higher-order product rule in your Assignment 1.

We now turn to Taylor's theorem for functions of several variables. We consider only scalar-valued functions for simplicity; the generalization to vector-valued functions is straightforward.

Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is of class C^k on a convex open set S. We can derive a Taylor expansion for $f(\mathbf{x})$ about a point $\mathbf{a} \in S$ by looking at the restriction of f to the line joining $\mathbf{a}$ and $\mathbf{x}$. That is, we set $\mathbf{h} = \mathbf{x} - \mathbf{a}$ and

$$g(t) = f(\mathbf{a} + t(\mathbf{x} - \mathbf{a})) = f(\mathbf{a} + t\mathbf{h}).$$

By the chain rule,

$$g'(t) = \mathbf{h} \cdot \nabla f(\mathbf{a} + t\mathbf{h}),$$

and hence

$$g^{(j)}(t) = (\mathbf{h} \cdot \nabla)^j f(\mathbf{a} + t\mathbf{h}),$$

where the expression on the right denotes the result of applying the directional derivative

$$\mathbf{h} \cdot \nabla = h_1 \frac{\partial}{\partial x_1} + \cdots + h_n \frac{\partial}{\partial x_n} \tag{1}$$

j times to f. The Taylor formula for g with $a = 0$ and $h = 1$,

$$g(1) = \sum_0^k \frac{g^{(j)}(0)}{j!} 1^j + (\text{remainder}),$$

therefore yields

$$f(\mathbf{a}+\mathbf{h}) = \sum_0^k \frac{(\mathbf{h}\cdot\nabla)^j f(\mathbf{a})}{j!} + R_{\mathbf{a},k}(\mathbf{h}), \tag{2}$$

where formulas for $R_{\mathbf{a},k}(\mathbf{h})$ can be obtained from the Lagrange or integral formulas for remainders, applied to g.

It is usually preferable, however, to rewrite (2) and the accompanying formulas for the remainder so that the partial derivatives of f appear more explicitly. To do this, we apply the multinomial theorem to the expression (1) to get

$$(\mathbf{h}\cdot\nabla)^j = \sum_{|\alpha|=j} \frac{j!}{\alpha!}\mathbf{h}^\alpha \partial^\alpha.$$

Substituting this into (2) and the remainder formulas, we obtain the following:

Theorem 2 (Taylor's Theorem in Several Variables)**.** *Suppose* $f : \mathbb{R}^n \to \mathbb{R}$ *is of class* C^{k+1} *on an open convex set* S*. If* $\mathbf{a} \in S$ *and* $\mathbf{a}+\mathbf{h} \in S$*, then*

$$f(\mathbf{a}+\mathbf{h}) = \sum_{|\alpha|\le k} \frac{\partial^\alpha f(\mathbf{a})}{\alpha!}\mathbf{h}^\alpha + R_{\mathbf{a},k}(\mathbf{h}), \tag{3}$$

where the remainder is given in Lagrange's form by

$$R_{\mathbf{a},k}(\mathbf{h}) = \sum_{|\alpha|=k+1} \partial^\alpha f(\mathbf{a}+c\mathbf{h})\frac{\mathbf{h}^\alpha}{\alpha!} \text{ for some } c \in (0,1). \tag{4}$$

and in integral form by

$$R_{\mathbf{a},k}(\mathbf{h}) = (k+1)\sum_{|\alpha|=k+1} \frac{\mathbf{h}^\alpha}{\alpha!}\int_0^1 (1-t)^k \partial^\alpha f(\mathbf{a}+t\mathbf{h})\,dt. \tag{5}$$

This result bears a pleasing similarity to the single-variable formulas — a triumph for multi-index notation! It may be reassuring, however, to see the formula for the second-order Taylor polynomial written out in the more familiar notation:

$$P_{\mathbf{a},2}(\mathbf{h}) = f(\mathbf{a}) + \sum_{j=1}^n \partial_j f(\mathbf{a})h_j + \frac{1}{2}\sum_{j,k=1}^n \partial_j\partial_k f(\mathbf{a})h_j h_k \tag{6}$$

$$= f(\mathbf{a}) + \sum_1^n \partial_j f(\mathbf{a})h_j + \frac{1}{2}\sum_{j=1}^n \partial_j^2 f(\mathbf{a})h_j^2 + \sum_{1\le j<k\le n} \partial_j\partial_k f(\mathbf{a})h_j h_k. \tag{7}$$

The first of these formulas is (2) with $k = 2$; the second one is (3). (Every multi-index α of order 2 is either of the form $(\ldots,2,\ldots)$ or $(\ldots,1,\ldots,1,\ldots)$, where the dots denote zero entries, so the sum over $|\alpha| = 2$ in (3) breaks up into the last two sums in (7).) Notice that the mixed derivatives $\partial_j\partial_k$ ($j \neq k$) occur twice in (6) (since $\partial_j\partial_k = \partial_k\partial_j$) but only once in (7) (since $j < k$ there); this accounts for the disappearance of the factor of $\frac{1}{2}$ in the last sum in (7).

As in the one-variable case, the following estimate for the remainder term follows from the Lagrange or integral formulas for it:

Corollary 1. *If f is of class C^{k+1} on S and $|\partial^\alpha f(\mathbf{x})| \leq M$ for $\mathbf{x} \in S$ and $|\alpha| = k+1$, then*

$$|R_{\mathbf{a},k}(\mathbf{h})| \leq \frac{M}{(k+1)!}\|\mathbf{h}\|^{k+1},$$

where

$$\|\mathbf{h}\| = |h_1| + |h_2| + \cdots + |h_n|.$$

Proof. It follows easily from either (5) or (4) that

$$|R_{\mathbf{a},k}(\mathbf{h})| \leq M \sum_{|\alpha|=k+1} \frac{|\mathbf{h}^\alpha|}{\alpha!},$$

and this last expression equals $M\|\mathbf{h}\|^{k+1}/(k+1)!$ by the multinomial theorem. ∎

As in the one-variable case, the Taylor polynomial $\sum_{|\alpha|\leq k}(\partial^\alpha f(\mathbf{a})/\alpha!)(\mathbf{x}-\mathbf{a})^\alpha$ is the *only* polynomial of degree $\leq k$ that agrees with $f(\mathbf{x})$ to order k at $\mathbf{x}-\mathbf{a}$, so the same algebraic devices are available to derive Taylor expansions of complicated functions from Taylor expansions of simpler ones.

EXAMPLE. Find the 3rd-order Taylor polynomial of $f(x,y) = e^{x^2+y}$ about $(x,y) = (0,0)$.

Solution. The direct method is to calculate all the partial derivatives of f of order ≤ 3 and plug the results into (3), but only a masochist would do this. Instead, use the familiar expansion for the exponential function, neglecting all terms of order higher than 3:

$$\begin{aligned} e^{x^2+y} &= 1 + (x^2+y) + \tfrac{1}{2}(x^2+y)^2 + \tfrac{1}{6}(x^2+y)^3 + (\text{order} > 3) \\ &= 1 + x^2 + y + \tfrac{1}{2}(x^4 + 2x^2y + y^2) + \tfrac{1}{6}(x^6 + 3x^4y + 3x^2y^2 + y^3) \\ &\qquad + (\text{order} > 3) \\ &= 1 + y + x^2 + \tfrac{1}{2}y^2 + x^2y + \tfrac{1}{6}y^3 + (\text{order} > 3). \end{aligned}$$

In the last line we have thrown the terms x^4, x^6, x^4y, and x^2y^2 into the garbage pail, since they are themselves of order > 3. Thus the answer is $1+y+x^2+\frac{1}{2}y^2+x^2y+\frac{1}{6}y^3$. Alternatively,

$$\begin{aligned} e^{x^2+y} = e^{x^2}e^y &= (1 + x^2 + \cdots)(1 + y + \tfrac{1}{2}y^2 + \tfrac{1}{6}y^3 + \cdots) \\ &= 1 + y + x^2 + \tfrac{1}{2}y^2 + x^2y + \tfrac{1}{6}y^3 + \cdots \end{aligned}$$

where the dots indicate terms of order > 3.