A note on multiindices

Ulrik Skre Fjordholm

February 18, 2021

Multiindices are an elegant way of taking derivatives of multivariate functions (functions depending on several real variables) and, in particular, Taylor expanding such functions. This note gives an introduction to multiindices and some things you can do with them, such as Taylor expanding multivariate functions and computing multivariate power series.

1 Multiindices

Let $d \in \mathbb{N}$. When we say that a function $f : \mathbb{R}^d \to \mathbb{R}$ is *smooth*, we mean that it is differentiable as many times as we require.

Notation. The sets of positive integers is $\mathbb{N} = \{1, 2, 3, \dots\}$, and the set of non-negative integers is $\mathbb{N}_0 = \{0, 1, 2, \dots\}$.

Definition. Let $d \in \mathbb{N}$.

- A *multiindex* is a vector $\alpha \in {\mathbb{N}_0}^d$, that is, a vector $\alpha = (\alpha_1, \dots, \alpha_d)$ where $\alpha_1, \dots, \alpha_d \in \mathbb{N}_0$. If α and β are multiindices then we write $\beta \leqslant \alpha$ if $\beta_i \leqslant \alpha_i$ for all $i = 1, \dots, d$.
- The *factorial* and *length* of a multiindex α are the numbers $\alpha! = \alpha_1!\alpha_2!\dots\alpha_d!$ and $|\alpha| = \alpha_1 + \alpha_2 + \dots + \alpha_d$.
- The *multinomial coefficients* are the numbers $\binom{|\alpha|}{\alpha}$, where

  $$\binom{|\alpha|}{\alpha} = \frac{|\alpha|!}{\alpha!} = \frac{(\alpha_1 + \dots + \alpha_d)!}{\alpha_1! \cdots \alpha_d!}.$$

  (We do not give meaning to $\binom{k}{\alpha}$ for $k \neq |\alpha|$.)
- If $x = (x_1, \dots, x_d) \in \mathbb{R}^d$ then $x^\alpha = x_1^{\alpha_1} \cdots x_d^{\alpha_d}$ (with the convention that $a^0 = 1$ for any $a \in \mathbb{R}$).
- For a smooth function $f : \mathbb{R}^d \to \mathbb{R}$ and a multiindex $\alpha \in {\mathbb{N}_0}^d$, we define the α-th partial derivative of f as

  $$f^{(\alpha)}(x) = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}}(x),$$

  that is, the partial derivative of f taken α_1 times with respect to x_1, α_2 times with respect to x_2, and so on. Note that *any* partial derivative can be written in terms of multiindices.

- A *multivariate polynomial* is a linear combination of products of terms of the form x_i^n for $n \in \mathbb{N}$ and $i \in \{1, \dots, d\}$. Thus, every multivariate polynomial p can be written as

$$p(x) = \sum_{|\alpha| \leqslant k} c_\alpha x^\alpha,$$

  where $c_\alpha \in \mathbb{R}$ are fixed coefficients and k is the *order* of the multivariate polynomial.

Example. If $f : \mathbb{R}^3 \to \mathbb{R}$ is smooth then $f^{((2,0,1))} = \frac{\partial^3 f}{\partial x_1^2 \partial x_3}$. See also the problems at the end of this note.

Remark. Let α be a multiindex. We can make the following combinatorial interpretations:

- Assume that we have distributed $|\alpha|$ objects into d different containers, where container i contains α_i objects ($i = 1, \dots, d$), and the objects are lying side-by-side (so they are ordered from left to right). Then $\alpha!$ is the number of distinct ways in which we can order these elements, without moving objects from one container to another.
- Given $|\alpha|$ objects and d containers, the multinomial coefficient $\binom{|\alpha|}{\alpha}$ is the number of ways in which we can place the $|\alpha|$ different objects into the d containers, while making sure that container i contains exactly α_i objects ($i = 1, \dots, d$).

The following results are examples of the elegance and simplicity that can be achieved in using multiindices.

Theorem (The Multinomial Theorem)**.** *Let $x \in \mathbb{R}^d$ and let $k \in \mathbb{N}$. Then*

$$(x_1 + x_2 + \dots + x_d)^k = \sum_{|\alpha|=k} \binom{|\alpha|}{\alpha} x^\alpha$$

where the sum is taken over all multiindices α of length $|\alpha| = k$.

Idea of proof. Use induction on d. You will need the binomial theorem, $(a+b)^k = \sum_{i=0}^k \binom{i}{k} a^i b^{k-i}$. Try writing this statement using multiindices of length $d = 2$. $\square$

Theorem (Leibniz' formula)**.** *Let $f, g : \mathbb{R}^d \to \mathbb{R}$ be smooth functions and let $\alpha \in {\mathbb{N}_0}^d$ be a multiindex. Then*

$$(fg)^{(\alpha)} = \sum_{\beta \leqslant \alpha} \binom{\alpha}{\beta} f^{(\beta)} g^{(\alpha-\beta)},$$

where the sum is taken over all multiindices $\beta \leqslant \alpha$, and we write $\binom{\alpha}{\beta} = \frac{\alpha!}{\beta!(\alpha-\beta)!}$.

Idea of proof. Use induction on d. You will need the standard Leibniz formula for functions of one variable, $(fg)^{(k)} = \sum_{i=0}^k \binom{i}{k} f^{(i)} g^{(k-i)}$. In the induction step $d \rightsquigarrow d+1$, write

$$(fg)^{(\alpha)} = \frac{\partial^{\alpha_{d+1}}}{\partial x_{d+1}^{\alpha_{d+1}}} (fg)^{(\bar{\alpha})},$$

where $\alpha = (\alpha_1, \dots, \alpha_d, \alpha_{d+1})$ and $\bar{\alpha} = (\alpha_1, \dots, \alpha_d)$. $\square$

Theorem (Taylor's formula)**.** *Let $f : \mathbb{R}^d \to \mathbb{R}$ be a smooth function, let $k \in \mathbb{N}$ and $z \in \mathbb{R}^d$. Then*

$$f(x) = \sum_{|\alpha| \leqslant k} \frac{f^{(\alpha)}(z)}{\alpha!}(x-z)^\alpha + \mathscr{R}_k(x)$$

where the sum is taken over all multiindices α of length $|\alpha| \leqslant k$, and the remainder term $\mathscr{R}_k$ satisfies

$$|\mathscr{R}_k(x)| \leqslant C\|x-z\|^{k+1}$$

for some $C > 0$.

Remark. The multivariate polynomial $p(x) = \sum_{|\alpha| \leqslant k} \frac{f^{(\alpha)}(z)}{\alpha!}(x-z)^\alpha$ is the k-th order Taylor expansion of f around z.

Proof of Taylor's formula. Fix a point $x \in \mathbb{R}^d$. Define the function

$$g(t) = f\big(z + t(x-z)\big) \qquad \forall\, t \in \mathbb{R}.$$

Then $g : \mathbb{R} \to \mathbb{R}$ is smooth and satisfies $g(0) = f(z)$, $g(1) = f(x)$. By Taylor's formula we know that

$$g(t) = \sum_{m=0}^{k} \frac{g^{(m)}(0)}{m!} t^m + R(t)$$

where $R(t) = \frac{g^{(k+1)}(s)}{(k+1)!} s^{k+1}$ for some s between 0 and t. We compute the derivatives of g using the chain rule:

$$\begin{aligned}
g'(t) &= \sum_{i=1}^{d} \frac{\partial f}{\partial x_i}\big(z + t(x-z)\big)(x_i - z_i), \\
g''(t) &= \sum_{i=1}^{d}\sum_{j=1}^{d} \frac{\partial^2 f}{\partial x_i \partial x_j}\big(z + t(x-z)\big)(x_i - z_i)(x_j - z_j) \\
&\ \vdots \\
g^{(m)}(t) &= \sum_{i_1=1}^{d} \cdots \sum_{i_m=1}^{d} \frac{\partial^m f}{\partial x_{i_1} \cdots \partial x_{i_m}}\big(z + t(x-z)\big)(x_{i_1} - z_{i_1}) \cdots (x_{i_m} - z_{i_m}).
\end{aligned}$$

Each term $\frac{\partial^m f}{\partial x_{i_1} \cdots \partial x_{i_m}}$ is of the form $f^{(\alpha)}$ for some multiindex α of length $|\alpha| = m$, and the corresponding coefficient $(x_{i_1} - z_{i_1}) \cdots (x_{i_m} - z_{i_m})$ can be written as $(x-z)^\alpha$. The number of different ways in which $f^{(\alpha)}$ can occur when written in the above form is precisely

$$\binom{m}{\alpha} = \frac{m!}{\alpha!}$$

(see the remark on page 2). Hence,

$$g^{(m)}(t) = \sum_{|\alpha| = m} \binom{m}{\alpha} f^{(\alpha)}\big(z + t(x-z)\big)(x-z)^\alpha.$$

Using the Taylor expansion of g we obtain

$$\begin{aligned}
f(x) = g(1) &= \sum_{m=0}^{k} \frac{g^{(m)}(0)}{m!} 1^m + R(1) \\
&= \sum_{m=0}^{k} \frac{1}{m!} \sum_{|\alpha|=m} \frac{m!}{\alpha!} f^{(\alpha)}\big(z + 0(x-z)\big)(x-z)^\alpha + R(1) \\
&= \sum_{m=0}^{k} \sum_{|\alpha|=m} \frac{1}{\alpha!} f^{(\alpha)}(z)(x-z)^\alpha + R(1) \\
&= \sum_{|\alpha|\leqslant k} \frac{1}{\alpha!} f^{(\alpha)}(z)(x-z)^\alpha + R(1).
\end{aligned}$$

We finally estimate $R(1)$: For some $s \in (0,1)$ we have

$$\begin{aligned}
|R(1)| &= \left| \frac{g^{(k+1)}(s)}{(k+1)!} s^{k+1} \right| \\
&= \frac{1}{(k+1)!} s^{k+1} \left| \sum_{|\alpha|=k+1} \frac{(k+1)!}{\alpha!} f^{(\alpha)}\big(z + s(x-z)\big)(x-z)^\alpha \right|
\end{aligned}$$

(as $|s| \leqslant 1$)

$$\leqslant \left| \sum_{|\alpha|=k+1} \frac{1}{\alpha!} f^{(\alpha)}\big(z + s(x-z)\big)(x-z)^\alpha \right|$$

(triangle inequality)

$$\leqslant \sum_{|\alpha|=k+1} \frac{1}{\alpha!} \left| f^{(\alpha)}\big(z + s(x-z)\big) \right| \left|(x-z)^\alpha\right|.$$

Let $\widetilde{C} > 0$ be a number that bounds all $(k+1)$th order derivatives of f, that is, $\left|f^{(\alpha)}(y)\right| \leqslant \widetilde{C}$ for all multiindices $|\alpha| = k+1$ and all y. Next, we have $|(x-z)^\alpha| = |x_1 - z_1|^{\alpha_1} \cdots |x_d - z_d|^{\alpha_d} \leqslant \|x-z\|^{\alpha_1} \cdots \|x-z\|^{\alpha_d} = \|x-z\|^{\alpha_1+\cdots+\alpha_d} = \|x-z\|^{k+1}$. Hence,

$$|R(1)| \leqslant \sum_{|\alpha|=k+1} \frac{1}{\alpha!} \widetilde{C} \|x-z\|^{k+1} = \widetilde{C} \|x-z\|^{k+1} \underbrace{\sum_{|\alpha|=k+1} \frac{1}{\alpha!}}_{=D} = C\|x-z\|^{k+1},$$

where $C = \widetilde{C}D$. (See also Problem 5.) □

2 Multivariate series

Just as for univariate power series,

$$\sum_{n=0}^{\infty} c_n (x-a)^n, \qquad x \in \mathbb{R}$$

we can develop a theory of multivariate power series

$$\sum_{\alpha} c_\alpha (x-a)^\alpha, \qquad x \in \mathbb{R}^d \tag{1}$$

where the "sum" is taken over all multiindices $\alpha \in {\mathbb{N}_0}^d$, $a \in \mathbb{R}^d$ is a fixed point and $c_\alpha \in \mathbb{R}$ are fixed coefficients. We say that the power series (1) *converges* at a point $x \in \mathbb{R}^d$ if the sequence of partial sums $\{S_n\}_{n\in\mathbb{N}}$ converges, where

$$S_n = \sum_{|\alpha| \leqslant n} c_\alpha (x-a)^\alpha.$$

(This is just a sequence of real numbers, and convergence of this sequence is meant in the usual sense.) For a power series (1), we define its *radius of convergence* as

$$R = \frac{1}{\limsup_{|\alpha|\to\infty} |c_\alpha|^{1/|\alpha|}}$$

(with the usual convention $1/0 = \infty$ and $1/\infty = 0$). The "lim sup" should be understood as

$$\limsup_{|\alpha|\to\infty} |c_\alpha|^{1/|\alpha|} = \lim_{k\to\infty} \left(\sup_{|\alpha|\geqslant k} |c_\alpha|^{1/|\alpha|} \right),$$

where the supremum is taken over all multiindices α of length $|\alpha| \geqslant k$.

Theorem (The Cauchy–Hadamard theorem)**.** *Let R be the radius of convergence of the multivariate power series* (1)*. Then* (1) *converges at all points x in the "rectangle"*

$$C(a;R) = (a_1 - R, a_1 + R) \times \cdots \times (a_d - R, a_d + R).$$

Proof. The proof follows the univariate version closely. If $R = 0$ then $C(a;R)$ is empty, so there is nothing to prove. If $R > 0$ then $\eta = \limsup_{|\alpha|\to\infty} |c_\alpha|^{1/|\alpha|}$ is finite, so for any $\varepsilon > 0$ there is some $N \in \mathbb{N}$ such that

$$\left| \eta - \sup_{|\alpha|\geqslant k} |c_\alpha|^{1/|\alpha|} \right| < \varepsilon \qquad \forall\ k \geqslant N.$$

In particular, $\sup_{|\alpha|\geqslant k} |c_\alpha|^{1/|\alpha|} < \varepsilon + \eta$. Let now $r \in (0,R)$, fix a point $x \in C(a;r)$ and let $\varepsilon = \frac{1}{2}(r^{-1} - \eta) > 0$. If $|\alpha| \geqslant N$ then

$$\begin{aligned} \bigl|c_\alpha (x-a)^\alpha\bigr| &= \Bigl(|c_\alpha|^{1/|\alpha|} |x_1 - a_1|^{\alpha_1/|\alpha|} \cdots |x_d - a_d|^{\alpha_d/|\alpha|} \Bigr)^{|\alpha|} \\ &\leqslant \Bigl((\eta+\varepsilon) r^{\alpha_1/|\alpha|} \cdots r^{\alpha_d/|\alpha|} \Bigr)^{|\alpha|} \\ &= \bigl((\eta+\varepsilon)r\bigr)^{|\alpha|} \\ &= \gamma^{|\alpha|}, \end{aligned} \tag{2}$$

where $\gamma = (\eta+\varepsilon)r = \frac{1+r\eta}{2}$. Since $r < R = \frac{1}{\eta}$ we have $r\eta < 1$ and hence $\gamma < 1$.

Let us write our series as

$$\sum_{k=0}^{\infty} \sum_{|\alpha|=k} c_\alpha (x-a)^\alpha.$$

If $k \geqslant N$ then we can bound

$$\left|\sum_{|\alpha|=k} c_\alpha (x-a)^\alpha\right| \leqslant \sum_{|\alpha|=k} \left|c_\alpha (x-a)^\alpha\right| \leqslant \sum_{|\alpha|=k} \gamma^{|\alpha|} = \binom{k+d-1}{d-1}\gamma^k,$$

since $\binom{k+d-1}{d-1}$ is the number of different multiindices $\alpha \in \mathbb{N}_0{}^d$ of length $|\alpha| = k$. The series

$$\sum_{k=N}^{\infty} \binom{k+d-1}{d-1}\gamma^k$$

is convergent (this can be seen by observing that $\binom{k+d-1}{d-1} \leqslant (k+d-1)^d$, and that $\sum_k (k+d-1)^d \gamma^k$ is convergent when $\gamma < 1$). Hence, by Weierstrass' M-test, we conclude that the series (1) converges uniformly in $C(a;r)$. $\square$

Remark. Unlike in the univariate case $d = 1$, we cannot state precisely for which x the series (1) diverges. In the computation (2), if for instance $|x_1 - a_1|$ is very small, but the other terms $|x_i - a_i|$ are large, then $|c_\alpha (x-a)^\alpha|$ is small, even though x is far away from $C(x;R)$.

Problems

1. Write a list of all multiindices $\alpha \in \mathbb{N}_0{}^2$ of length $|\alpha| \leqslant 2$.

2. Write out the expression

$$\sum_{|\alpha|=2} \alpha! x^\alpha$$

where $x = (a,b)$ is some point in $\mathbb{R}^2$.

(Here and elsewhere we use the convention that α denotes a multiindex, so the sum runs over all pairs of nonnegative integers $\alpha = (\alpha_1, \alpha_2) \in \mathbb{N}_0{}^2$ whose sum $|\alpha| = \alpha_1 + \alpha_2$ equals 2.)

3. Compute the partial derivative $f^{(\alpha)}$ for all multiindices α of length $|\alpha| = 1$ and $|\alpha| = 2$, for each of the following functions:

(a) $f(x) = \sin(x_1 x_2^2 - x_3)$ for $x = (x_1, x_2, x_3) \in \mathbb{R}^3$

(b) $f(x) = \|x\|^2$ for $x \in \mathbb{R}^n$

(c) $f(x) = \|x\|$ for $x \in \mathbb{R}^n$

4. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by $f(x) = \sin(x_1 - x_2)e^{1+x_2}$.

(a) Compute the first order Taylor expansion around $z = (0,0)$ of the function f.

(b) Write f as a power series by letting $k \to \infty$ in Taylor's formula. (You will need to find a general expression for $f^{(\alpha)}(z)$.)

5. Use the Multinomial Theorem to prove that

$$\sum_{|\alpha|=k} \frac{1}{\alpha!} = \frac{d^k}{k!}.$$

(*Hint: Write* $d = 1 + 1 + \cdots + 1$.)