DIFFERENTIAL GEOMETRY

Taylor’s formula in \( \mathbb{R}^n \)

We will need Taylor’s formula for a function of several variables. Since it is a consequence of the one variable formula, I start with that one.

Let \( I \) be an open interval in \( \mathbb{R} \) and let \( f : I \to \mathbb{R} \) be \( m \) times differentiable, where \( m \geq 1 \). Let \( a \in I \). Then

\[
f(t) = f(a) + \int_a^t f'(s) \, ds
\]

(1)

for all \( t \in I \).

This is actually the first step in Taylor’s formula; as far as we can go if \( m = 1 \). But if \( m \geq 2 \), we can write

\[
f'(s) = f'(a) + \int_a^s f''(\sigma) \, d\sigma,
\]

hence

\[
f(t) = f(a) + \int_a^t \left( f'(a) + \int_a^s f''(\sigma) \, d\sigma \right) \, ds = f(a) + f'(a) \int_a^t \, ds + \int_a^t \int_a^s f''(\sigma) \, d\sigma \, ds
\]

(We changed the order of integration in the double integral.) Calling the variable of integration again \( s \) for all \( t \), we now have:

\[
f(t) = f(a) + f'(a)(t - a) + \int_a^t (t - s) f''(s) \, ds
\]

(2)

for all \( t \in I \). If \( m \geq 3 \) we can repeat the process:

\[
f(t) = f(a) + f'(a)(t - a) + \int_a^t (t - s) \left( f''(a) + \int_a^s f'''(\sigma) \, d\sigma \right) \, ds
\]

\[
= f(a) + f'(a)(t - a) + f''(a) \int_a^t (t - s) \, ds + \int_a^t \int_a^s (t - \sigma) f'''(\sigma) \, d\sigma \, ds
\]

\[
= f(a) + f'(a)(t - a) + \frac{1}{2} f''(a)(t - a)^2 + \int_a^t \left( \int_\sigma (t - \sigma) \, ds \right) f'''(\sigma) \, d\sigma
\]

and performing the inner integral in the last expression, replacing again \( \sigma \) by \( s \), we get

\[
f(t) = f(a) + f'(a)(t - a) + \frac{1}{2} f''(a)(t - a)^2 + \frac{1}{2} \int_a^t (t - s)^2 f'''(s) \, ds.
\]

(3)

for all \( t \in I \). A pattern is emerging. It is now easy to prove by induction:

**Theorem 1** Assume \( I \) is an open interval in \( \mathbb{R} \), \( m \in \mathbb{N} \cup \{0\} \), and let \( f : I \to \mathbb{R} \) be \( m + 1 \) times continuously differentiable. Then

\[
f(t) = \sum_{k=0}^m \frac{1}{k!} f^{(k)}(a)(t - a)^k + \frac{1}{m!} \int_a^t (t - s)^m f^{(m+1)}(s) \, ds.
\]

(4)

This is Taylor’s Theorem with the remainder in integral form. In this context, \( P_m(t) = \sum_{k=0}^m \frac{1}{k!} f^{(k)}(a)(t - a)^k \) is called the \( m \)-th Taylor polynomial and \( R_m(t) = \frac{1}{m!} \int_a^t (t - s)^m f^{(m)}(s) \, ds \) is the remainder. To get the differential form of the remainder, one can use the following integral mean value theorem:
Theorem 2 Let \( f, g : [a, b] \to \mathbb{R} \) be continuous functions; assume \( g(s) \geq 0 \) for all \( s \in [a, b] \). There exists \( c \in (a, b) \) such that

\[
\int_a^b f(s)g(s) \, ds = f(c) \int_a^b g(s) \, ds.
\]

Proving this theorem should be a relatively simple exercise for a second year graduate student. I will limit myself to observe that the theorem is also true if \( g(s) \leq 0 \) for all \( s \in [a, b] \). In fact, then there is \( c \) so

\[
\int_a^b f(s)g(s) \, ds = -\int_a^b f(s)(-g(s)) \, ds = -f(c) \int_a^b (-g(s)) \, ds = f(c) \int_a^b g(s) \, ds.
\]

In writing out Taylor’s formula one doesn’t have to assume \( t \geq a \), just \( t \in I \). In (4), the variable of integration \( s \) is between \( a \) and \( t \), thus \( t - s \) is either always non-negative or always non-positive; the same holds for \((t - s)^m\) (with the additional fact that it is always non-negative if \( m \) is even). Thus there is \( c \) between \( a \) and \( t \) such that

\[
\int_a^t (t - s)^m f^{(m+1)}(s) \, ds = f^{(m+1)}(c) \int_a^t (t - s)^m \, ds = \frac{1}{m+1} f^{(m+1)}(c)(t - a)^{m+1}.
\]

We thus get Schlömilch’s form of the remainder: There is \( c \) between \( a \) and \( t \) such that

\[
R_m(t) = \frac{1}{(m+1)!} f^{(m+1)}(c)(t - a)^{m+1}.
\] (5)

The number \( c \) depends, of course, on \( t \) and on \( m \). If \( t = a \), then \( c = t = a \).

The \( n \)-dimensional case. Assume now \( U \) is an open convex subset of \( \mathbb{R}^n \) (for example an open ball or an open box), \( f : U \to \mathbb{R} \) an \( m + 1 \)-times differentiable function, \( a = (a_1, \ldots, a_n) \in U \). For \( x \in U \) (to be kept fixed for a while) we can then define a one variable function \( F \) by

\[
F(t) = f(a + t(x - a));
\]

because \( U \) is open and convex this function is defined (at least) in an open interval \( I \supset [0, 1] \). Applying (4) with \( f \) replaced by \( F \), \( a \) by \( 0 \), and \( t \) by 1 (so the factors of \((t - a)^k\) appearing in that formula are all equal to 1), we get

\[
f(x) = F(1) = \sum_{k=0}^m \frac{1}{k!} F^{(k)}(0) + \frac{1}{m!} \int_0^1 (1 - s)^m F^{(m+1)}(s) \, ds.
\]

But now we need to calculate the derivatives of \( F \). A certain notation can be useful here. A multi-index is an \( n \)-tuple of non-negative integers. We’ll use mostly Greek letters from the beginning of the alphabet to denote them. To repeat, a multi-index is an \( n \)-tuple \( \alpha = (\alpha_1, \ldots, \alpha_n) \) such that \( \alpha_1, \ldots, \alpha_n \) are non-negative integers. If \( x = (x_1, \ldots, x_n) \in \mathbb{R}^n \) and \( \alpha = (\alpha_1, \ldots, \alpha_n) \) is a multi-index, we define

\[
x^{\alpha} = x_1^{\alpha_1} \cdots x_n^{\alpha_n}.
\]

For example, if \( n = 2 \), \( \alpha = (1, 3) \), then \( x^{\alpha} = x_1 x_2^3 \); if \( n = 5 \) and \( \alpha = (2, 3, 0, 1, 4) \), then \( x^{\alpha} = x_1^2 x_2^3 x_3 x_4^4 \). If \( \alpha = (0, \ldots, 0) \) is an \( n \)-tuple of 0’s (we write \( \alpha = 0 \)), then \( x^{\alpha} = 1 \) for all \( x \in \mathbb{R}^n \), even if some component of \( x \) is zero. That is \( 0^0 = 1 \) in this context. If \( \alpha = (\alpha_1, \ldots, \alpha_n) \) is a multi-index, the length of \( \alpha \), denoted by \( |\alpha| \) is defined by \( |\alpha| = \alpha_1 + \cdots + \alpha_n \). There is rarely if ever any danger of confusing the length of \( \alpha \) as a multi-index with its Euclidean norm as an element of \( \mathbb{R}^n \); which is which should be clear from the context. If \( f \) is a function of \( n \) variables, one also uses multi-indices to simplify the notation of higher order derivatives, in the continuously differentiable situation when the order of differentiation is irrelevant. Thus one defines for \( \alpha = (\alpha_1, \ldots, \alpha_n) \) a multi-index:

\[
D^{\alpha} f = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}.
\]

For example:

- If \( n = 2 \), \( \alpha = (1, 3) \), then \( D^{\alpha} f = \frac{\partial^4}{\partial x_1 \partial x_2^3} \).
If \( n = 5 \), \( \alpha = (2, 3, 0, 1, 4) \), then \( D^\alpha f = \frac{\partial^{10} f}{\partial x_1^2 \partial x_2^3 \partial x_4 \partial x_5^4} \),

If \( \alpha = 0 \) (the 0 multi-index) then \( D^0 f = f \),

For general \( n \) one should notice that surprisingly enough

\[
\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j} \neq \sum_{|\alpha|=2} D^\alpha f.
\]

This is because in the sum of the left hand side terms with \( i \neq j \) appear twice, on the right hand side only once. For example, if \( n = 2 \), then the sum of the left hand side is

\[
\frac{\partial^2 f}{\partial x_1^2} + \frac{\partial^2 f}{\partial x_1 \partial x_2} + \frac{\partial^2 f}{\partial x_2 \partial x_1} + \frac{\partial^2 f}{\partial x_2^2} = \frac{\partial^2 f}{\partial x_1^2} + 2 \frac{\partial^2 f}{\partial x_1 \partial x_2} + \frac{\partial^2 f}{\partial x_2^2},
\]

while the sum on the right hand side equals

\[
D^{(2,0)} f + D^{(1,1)} f + D^{(0,2)} f = \frac{\partial^2 f}{\partial x_1^2} + \frac{\partial^2 f}{\partial x_1 \partial x_2} + \frac{\partial^2 f}{\partial x_2^2}.
\]

Considering that if the order of derivation is 3 a term like \( \frac{\partial^3 f}{\partial x_1 \partial x_2 \partial x_3} \) appears not three times but six times in a sum involving all derivatives of order 3, we define

\[
\alpha! = \alpha_1! \cdots \alpha_n!.
\]

if \( \alpha = (\alpha_1, \ldots, \alpha_n) \) is a multi-index. Then we can write:

\[
\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j} = 2! \sum_{|\alpha|=2} \frac{1}{\alpha!} D^\alpha f.
\]

Of course, \( 2! = 2 \).

With the notation just developed, we can prove by induction:

\[
F^{(k)}(t) = k! \sum_{|\alpha|=k} \frac{1}{\alpha!} D^\alpha f(a + t(x-a))(x-a)^\alpha.
\]

Taylor’s theorem for the multivariable case follows:

**Theorem 3** Let \( f : U \to \mathbb{R} \) be \( m + 1 \)-times continuously differentiable, where \( U \) is an open convex subset of \( \mathbb{R}^n \), \( m \geq 0 \). Let \( a \in U \). Then for every \( x \in U \),

\[
f(x) = \sum_{|\alpha| \leq m} \frac{1}{\alpha!} D^\alpha f(a)(x-a)^\alpha + R_m(x),
\]

where

\[
R_m(x) = \sum_{|\alpha|=m+1} \frac{1}{\alpha!} D^\alpha f(\xi)(x-a)^\alpha,
\]

where \( \xi \) is a point on the line segment joining \( a \) to \( x \).

This is usually the most useful form. But occasionally we need the integral form for \( R_m \), which is

\[
R_m(x) = (m+1) \sum_{|\alpha|=m+1} \frac{(x-a)^\alpha}{\alpha!} \int_0^1 D^\alpha f(a + s(x-a)) \, ds.
\]

This is because \( x \mapsto \int_0^1 D^\alpha f(a + s(x-a)) \, ds \) is easily seen to be a continuous function on \( U \) and we get the following
**Corollary 4** Let $f : U \to \mathbb{R}$ be $m+1$-times continuously differentiable, where $U$ is an open convex subset of $\mathbb{R}^n$, $m \geq 0$. Let $a \in U$. Then

$$f(x) = \sum_{|\alpha| \leq m} \frac{(x-a)^\alpha}{\alpha!} D^\alpha f(a) + \sum_{|\alpha| = m+1} (x-a)^\alpha g_\alpha(x),$$

where $g_\alpha : U \to \mathbb{R}$ is continuous for each multi-index $\alpha$ of length $m+1$.

The functions $g_\alpha$ depend also, of course, on $a$. For the time being we need only a very modest version of this result, the case $m = 0$. Notice that in the corollary in question the $m+1$-st derivative does not appear at all. But it is needed for the corollary to hold. Anyway, the case $m = 0$ of the Corollary states:

**Corollary 5** Let $f : U \to \mathbb{R}$ be continuously differentiable, where $U$ is an open convex subset of $\mathbb{R}^n$, $m \geq 0$. Let $a \in U$. Then

$$f(x) = f(a) + \sum_{j=1}^m (x_j - a_j) g_j(x),$$

where $g_j : U \to \mathbb{R}$ is continuous for $j = 1, \ldots, m$ and $g_j(a) = \frac{\partial f}{\partial x_j}(a)$.

**Proof.** According to (8),

$$g_j(x) = \int_0^1 \frac{\partial f}{\partial x_j}(a + s(x-a)) \, ds, \quad j = 1, \ldots, n;$$

the result follows taking $x = a$. □