11. Higher derivatives

We first record a very useful:

**Theorem 11.1.** Let $A \subset \mathbb{R}^n$ be an open subset. Let $f: A \rightarrow \mathbb{R}^m$ and $g: A \rightarrow \mathbb{R}^m$ be two functions and suppose that $P \in A$. Let $\lambda \in A$ be a scalar.

If $f$ and $g$ are differentiable at $P$, then

1. $f + g$ is differentiable at $P$ and $D(f + g)(P) = Df(P) + Dg(P)$.
2. $\lambda \cdot f$ is differentiable at $P$ and $D(\lambda f)(P) = \lambda D(f)(P)$.

Now suppose that $m = 1$.

3. $fg$ is differentiable at $P$ and $D(fg)(P) = D(f)(P)g(P) + f(P)D(g)(P)$.
4. If $g(P) \neq 0$, then $fg$ is differentiable at $P$ and

$$D(f/g)(P) = \frac{D(f)(P)g(P) - f(P)D(g)(P)}{g^2(P)}.$$

If the partial derivatives of $f$ and $g$ exist and are continuous, then (11.1) follows from the well-known single variable case. One can prove the general case of (11.1), by hand (basically lots of $\varepsilon$’s and $\delta$’s). However, perhaps the best way to prove (11.1) is to use the chain rule, proved in the next section.

What about higher derivatives?

**Definition 11.2.** Let $A \subset \mathbb{R}^n$ be an open set and let $f: A \rightarrow \mathbb{R}$ be a function. The $k$th order partial derivative of $f$, with respect to the variables $x_{i_1}, x_{i_2}, \ldots x_{i_k}$ is the iterated derivative

$$\frac{\partial^k f}{\partial x_{i_k} \partial x_{i_{k-1}} \ldots \partial x_{i_2} \partial x_{i_1}}(P) = \frac{\partial}{\partial x_{i_k}} \left( \frac{\partial}{\partial x_{i_{k-1}}} \left( \ldots \frac{\partial f}{\partial x_{i_2}} \left( \frac{\partial f}{\partial x_{i_1}} \right) \ldots \right) \right)(P).$$

We will also use the notation $f_{x_{i_k}x_{i_{k-1}}\ldots x_{i_2}x_{i_1}}(P)$.

**Example 11.3.** Let $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ be the function $f(x, t) = e^{-at} \cos x$.

Then

$$f_{xx}(x, t) = \frac{\partial}{\partial x} \left( \frac{\partial}{\partial x} (e^{-at} \cos x) \right) = \frac{\partial}{\partial x} (-e^{-at} \sin x) = -e^{-at} \cos x.$$
On the other hand,
\[ f_{xt}(x, t) = \frac{\partial}{\partial x} \left( \frac{\partial}{\partial t} (e^{-at}\cos x) \right) \]
\[ = \frac{\partial}{\partial x} (-ae^{-at}\cos x) \]
\[ = ae^{-at}\sin x. \]

Similarly,
\[ f_{tx}(x, t) = \frac{\partial}{\partial t} \left( \frac{\partial}{\partial x} (e^{-at}\cos x) \right) \]
\[ = \frac{\partial}{\partial t} (-e^{-at}\sin x) \]
\[ = ae^{-at}\sin x. \]

Note that
\[ f_t(x, t) = -ae^{-at}\cos x. \]

It follows that \( f(x, t) \) is a solution to the \textbf{Heat equation}:
\[
a \frac{\partial^2 f}{\partial x^2} = \frac{\partial f}{\partial t}. \]

\textbf{Definition 11.4.} Let \( A \subset \mathbb{R}^n \) be an open subset and let \( f: A \rightarrow \mathbb{R}^m \) be a function. We say that \( f \) is of \textbf{class} \( C^k \) if all \( k \)th partial derivatives exist and are continuous.

We say that \( f \) is of \textbf{class} \( C^\infty \) (aka \textbf{smooth}) if \( f \) is of class \( C^k \) for all \( k \).

In lecture 10 we saw that if \( f \) is \( C^1 \), then it is differentiable.

\textbf{Theorem 11.5.} Let \( A \subset \mathbb{R}^n \) be an open subset and let \( f: A \rightarrow \mathbb{R}^m \) be a function.

If \( f \) is \( C^2 \), then
\[
\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i},
\]
for all \( 1 \leq i, j \leq n \).

The proof uses the Mean Value Theorem.

Suppose we are given \( A \subset \mathbb{R} \) an open subset and a function \( f: A \rightarrow \mathbb{R} \) of class \( C^1 \). The objective is to find a solution to the equation
\[ f(x) = 0.\]

Newton’s method proceeds as follows. Start with some \( x_0 \in A \). The best linear approximation to \( f(x) \) in a neighbourhood of \( x_0 \) is given by
\[
f(x_0) + \frac{f'(x_0)}{2}(x - x_0).\]
If \( f'(x_0) \neq 0 \), then the linear equation
\[
    f(x_0) + f'(x_0)(x - x_0) = 0,
\]
has the unique solution,
\[
    x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}.
\]
Now just keep going (assuming that \( f'(x_i) \) is never zero),
\[
    x_1 = x_0 - \frac{f(x_0)}{f'(x_0)},
\]
\[
    x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}
\]
\[
    \vdots
\]
\[
    x_n = x_{n-1} - \frac{f(x_{n-1})}{f'(x_{n-1})}.
\]

Claim 11.6. Suppose that \( x_\infty = \lim_{n \to \infty} x_n \) exists and \( f'(x_\infty) \neq 0 \). Then \( f(x_\infty) = 0 \).

Proof of [11.6]. Indeed, we have
\[
    x_n = x_{n-1} - \frac{f(x_{n-1})}{f'(x_{n-1})}.
\]
Take the limit as \( n \) goes to \( \infty \) of both sides:
\[
    x_\infty = x_\infty - \frac{f(x_\infty)}{f'(x_\infty)},
\]
we we used the fact that \( f \) and \( f' \) are continuous and \( f'(x_\infty) \neq 0 \). But then
\[
    f(x_\infty) = 0,
\]
as claimed. \( \square \)

Suppose that \( A \subset \mathbb{R}^n \) is open and \( f : A \to \mathbb{R}^n \) is a function. Suppose that \( f \) is \( C^1 \) (that is, suppose each of the coordinate functions \( f_1, f_2, \ldots, f_n \) is \( C^1 \)).

The objective is to find a solution to the equation
\[
    f(P) = \vec{0}.
\]

Start with any point \( P_0 \in A \). The best linear approximation to \( f \) at \( P_0 \) is given by
\[
    f(P_0) + Df(P_0)\overrightarrow{PP_0}.
\]
Assume that $Df(P_0)$ is an invertible matrix, that is, assume that $\det Df(P_0) \neq 0$. Then the inverse matrix $Df(P_0)^{-1}$ exists and the unique solution to the linear equation

$$f(P_0) + Df(P_0)\overrightarrow{PP_0} = \overrightarrow{0},$$

is given by

$$P_1 = P_0 - Df(P_0)^{-1}f(P_0).$$

Notice that matrix multiplication is not commutative, so that there is a difference between $Df(P_0)^{-1}f(P_0)$ and $f(P_0)Df(P_0)^{-1}$. If possible, we get a sequence of solutions,

$$P_1 = P_0 - Df(P_0)^{-1}f(P_0)
\quad P_2 = P_1 - Df(P_1)^{-1}f(P_1)
\quad \vdots
\quad P_n = P_{n-1} - Df(P_{n-1})^{-1}f(P_{n-1}).$$

Suppose that the limit $P_\infty = \lim_{n \to \infty} P_n$ exists and that $Df(P_\infty)$ is invertible. As before, if we take the limit of both sides, this implies that

$$f(P_\infty) = \overrightarrow{0}.$$

Let us try a concrete example.

**Example 11.7.** Solve

$$x^2 + y^2 = 1$$
$$y^2 = x^3.$$

First we write down an appropriate function, $f : \mathbb{R}^2 \to \mathbb{R}^2$, given by $f(x, y) = (x^2 + y^2 - 1, y^2 - x^3)$. Then we are looking for a point $P$ such that

$$f(P) = (0, 0).$$

Then

$$Df(P) = \begin{pmatrix} 2x & 2y \\ -3x^2 & 2y \end{pmatrix}.$$ 

The determinant of this matrix is

$$4xy + 6x^2y = 2xy(2 + 3x).$$

Now if we are given a $2 \times 2$ matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

suppose that $Df(P_\infty)$ is invertible.
then we may write down the inverse by hand,

\[ \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}. \]

So

\[ Df(P)^{-1} = \frac{1}{2xy(2 + 3x)} \begin{pmatrix} 2y & -2y \\ 3x^2 & 2x \end{pmatrix}. \]

So,

\[
Df(P)^{-1} f(P) = \frac{1}{2xy(2 + 3x)} \begin{pmatrix} 2y & -2y \\ 3x^2 & 2x \end{pmatrix} \begin{pmatrix} x^2 + y^2 - 1 \\ y^2 - x^3 \end{pmatrix} \\
= \frac{1}{2xy(2 + 3x)} \begin{pmatrix} 2x^2y - 2y + 2x^3y \\ x^4 + 3x^2y^2 - 3x^2 + 2xy^2 \end{pmatrix}.
\]

One nice thing about this method is that it is quite easy to implement on a computer. Here is what happens if we start with \((x_0, y_0) = (5, 2)\),

\[
(x_0, y_0) = (5.00000000000000, 2.00000000000000) \\
(x_1, y_1) = (3.24705882352941, -0.617647058823529) \\
(x_2, y_2) = (2.09875150983980, 1.37996311951634) \\
(x_3, y_3) = (1.37227480405610, 0.56122096705054) \\
(x_4, y_4) = (0.959201654346683, 0.50383950409063) \\
(x_5, y_5) = (0.787655203525685, 0.657830227357845) \\
(x_6, y_6) = (0.755918792660404, 0.655438554539110),
\]

and if we start with \((x_0, y_0) = (5, 5)\),

\[
(x_0, y_0) = (5.00000000000000, 5.00000000000000) \\
(x_1, y_1) = (3.24705882352941, 1.85294117647059) \\
(x_2, y_2) = (2.09875150983980, 0.363541705259258) \\
(x_3, y_3) = (1.37227480405610, -0.306989760884339) \\
(x_4, y_4) = (0.959201654346683, -0.561589294711320) \\
(x_5, y_5) = (0.787655203525685, -0.644964218428458) \\
(x_6, y_6) = (0.755918792660404, -0.655519172668858).
\]

One can sketch the two curves and check that these give reasonable solutions. One can also check that \((x_6, y_6)\) lie close to the two given curves, by computing \(x_6^2 + y_6^2 - 1\) and \(y_6^2 - x_6^3\).