Lecture 4

2.2 Conditions for Differentiability

In this lecture we will discuss conditions that guarantee differentiability. First, we begin with a review of important results from last lecture.

Let U be an open subset of $\mathbb{R}^n$, let $f : U \to \mathbb{R}^n$ be a map, and let $a \in U$.

We defined f to be differentiable at a if there exists a linear map $B : \mathbb{R}^n \to \mathbb{R}^m$ such that for $h \in \mathbb{R}^n - \{0\}$,

$$\frac{f(a+h) - f(a) - Bh}{|h|} \to 0 \text{ as } h \to 0. \tag{2.24}$$

If such a B exists, then it is unique and $B = Df(a)$. The matrix representing B is the Jacobian matrix $J_f(a) = \left[\frac{\partial f_i}{\partial x_j}(a)\right]$, where $f = (f_1, \ldots, f_m)$.

Note that the mere existence of all of the partial derivatives in the Jacobian matrix does not guarantee differentiability.

Now we discuss conditions that guarantee differentiability.

Theorem 2.7. *Suppose that all of the partial derivatives $\frac{\partial f_i}{\partial x_j}$ in the Jacobian matrix exist at all points $x \in U$, and that all of the partial derivatives are continuous at $x = a$. Then f is differentiable at a.*

Sketch of Proof. This theorem is very elegantly proved in Munkres, so we will simply give the general ideas behind the proof here.

First, we look at the case $n = 2, m = 1$. The main ingredient in the proof is the Mean Value Theorem from 1-D calculus, which we state here without proof.

Mean Value Theorem. *Given an interval $[a, b] \subseteq \mathbb{R}$ and a map $\phi : [a, b] \to \mathbb{R}$, if ϕ is continuous on $[a, b]$ and differentiable on (a, b), then there exists a point $c \in (a, b)$ such that $\phi(b) - \phi(a) = \phi'(c)(b - a)$.*

Now we continue with the proof. Let f be a map $f : U \to \mathbb{R}$, where $U \subseteq \mathbb{R}^2$. So, f is a function of two variables $f = f(x_1, x_2)$. Consider a point $a = (a_1, a_2) \in U$ and any point $h \in \mathbb{R}^2 - \{0\}$ "close" to zero, where by close we mean $a + h \in U$. We want to compute $f(a + h) - f(a)$.

$$\begin{aligned} f(a+h) - f(a) &= f(a_1 + h_1, a_2 + h_2) - f(a_1, a_2) \\ &= f(a_1 + h_1, a_2 + h_2) - f(a_1, a_2 + h_2) \\ &\quad + f(a_1, a_2 + h_2) - f(a_1, a_2). \end{aligned} \tag{2.25}$$

Thinking of the first two terms as functions of the first argument only, and thinking of the last two terms as functions of the second term only, and applying the Mean

Value Theorem to each pair of terms, we obtain

$$\begin{aligned} f(a+h)-f(a) &= \frac{\partial f}{\partial x_1}(c_1, a_2+h_2)h_1 \\ &+ \frac{\partial f}{\partial x_2}(a_1, d_2)h_2, \end{aligned} \tag{2.26}$$

where $a_1 < c_1 < a_1 + h_1$ and $a_2 < d_2 < a_2 + h_2$. This can be rewritten as

$$f(a+h)-f(a) = \frac{\partial f}{\partial x_1}(c)h_1 + \frac{\partial f}{\partial x_2}(d)h_2, \tag{2.27}$$

where $c = (c_1, a_2 + h_2)$ and $d = (a_1, d_2)$.

We want to show that $(f(a+h) - f(a) - Df(a)h)/|h| \to 0$ as $h \to 0$, where $Df(a) = \left[\frac{\partial f}{\partial x_1}(a), \frac{\partial f}{\partial x_2}(a)\right]$. Using our previously derived expression for $f(a+h)-f(a)$, we find that

$$\begin{aligned} f(a+h)-f(a)-Df(a)h &= f(a+h)-f(a)-\frac{\partial f}{\partial x_1}(a)h_1 - \frac{\partial f}{\partial x_2}(a)h_2 \\ &= \left(\frac{\partial f}{\partial x_1}(c) - \frac{\partial f}{\partial x_1}(a)\right)h_1 + \left(\frac{\partial f}{\partial x_2}(d) - \frac{\partial f}{\partial x_2}(a)\right)h_2. \end{aligned} \tag{2.28}$$

We can use the sup norm to show that

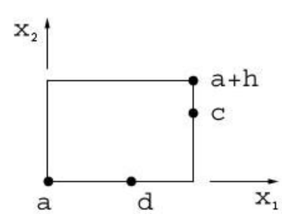
$$|f(a+h)-f(a)-Df(a)h| \le \left|\frac{\partial f}{\partial x_1}(c) - \frac{\partial f}{\partial x_1}(a)\right| |h_1| + \left|(\frac{\partial f}{\partial x_2}(d) - \frac{\partial f}{\partial x_2}(a)\right| |h_2|, \tag{2.29}$$

from which it follows that

$$\frac{|f(a+h)-f(a)-Df(a)h|}{|h|} \le \left|\frac{\partial f}{\partial x_1}(c) - \frac{\partial f}{\partial x_1}(a)\right| + \left|(\frac{\partial f}{\partial x_2}(d) - \frac{\partial f}{\partial x_2}(a)\right|, \tag{2.30}$$

where we used the fact that $|h| = \max(|h_1|, |h_2|)$.

Notice that as $h \to 0$, both $c \to a$ and $d \to a$, as can be easily seen using the following diagram. This means that the r.h.s. of Equation (2.30) goes to zero as h



goes to zero, because the partial derivatives are continuous. It follows that the l.h.s. goes to zero, which completes our proof.

The proof in n dimensions is similar to the above proof, but the details are harder to follow. $\square$

We now introduce a useful class of functions.

Definition 2.8. Given $U \subseteq \mathbb{R}^n$ and $f : U \to \mathbb{R}$, we define

$$f \in \mathcal{C}^1(U) \iff \frac{\partial f}{\partial x_i}, i = 1, \ldots, n \text{ exist and are continuous at all points } x \in U. \tag{2.31}$$

Similarly, we define

$$f \in \mathcal{C}^2(U) \iff \frac{\partial f}{\partial x_i} \in \mathcal{C}^1(U), i = a, \ldots, n. \tag{2.32}$$

$$f \in \mathcal{C}^k(U) \iff \frac{\partial f}{\partial x_i} \in \mathcal{C}^{k-1}(U), i = a, \ldots, n. \tag{2.33}$$

$$f \in \mathcal{C}^\infty(U) \iff f \in \mathcal{C}^k(U) \forall k. \tag{2.34}$$

If f is multiply differentiable, then you can perform higher order mixed partial derivatives.

One of the fundamental theorems of calculus is that the order of the partial derivatives can be taken in any order. For example,

$$\frac{\partial}{\partial x_i}\left(\frac{\partial}{\partial x_j}\right) = \frac{\partial}{\partial x_j}\left(\frac{\partial}{\partial x_i}\right) \equiv \frac{\partial^2 f}{\partial x_i \partial x_j} \tag{2.35}$$

Let's do the proof for this case. Let $U \subseteq \mathbb{R}^2$ and $f = f(x_1, x_2)$. We prove the following claim:

Claim.

$$\frac{\partial}{\partial x_i}\left(\frac{\partial}{\partial x_j}\right) = \frac{\partial}{\partial x_j}\left(\frac{\partial}{\partial x_i}\right). \tag{2.36}$$

Proof. Take $a \in U$ written in components as $a = (a_1, a_2)$, and take $h = (h_1, h_2) \in \mathbb{R}^2 - \{0\}$ such that $a + h \in U$. That is, take $h \approx 0$.

Define

$$\Delta(h) = f(a_1 + h_1, a_2 + h_2) - f(a_1, a_2 + h_2) - f(a_1 + h_1, a_2) + f(a_1, a_2), \tag{2.37}$$

and define

$$\phi(s) = f(a_1 + h_1, s) - f(a_1, s), \tag{2.38}$$

where $a_2 \leq s \leq a_2 + h_2$. We find that

$$\begin{aligned} \Delta(h) &= \phi(a_2 + h_2) - \phi(a_2) \\ &= \phi'(c_2)h_2, a_2 < c_2 < a_2 + h_2, \end{aligned} \tag{2.39}$$

by the Mean Value Theorem. Writing out ϕ' using partial derivatives of f, and using the Mean Value Theorem again, we find

$$\begin{aligned}
\Delta(h) &= \left(\frac{\partial f}{\partial x_2}(a_1 + h_1, c_2) - \frac{\partial f}{\partial x_1}(a_1, c_2)\right) h_2 \\
&= \left(\frac{\partial}{\partial x_1}\left(\frac{\partial f}{\partial x_2}(c_1, c_2) h_1\right)\right) h_2, a_1 < c_1 < a_1 + h_1 \\
&= \left(\frac{\partial}{\partial x_1}\left(\frac{\partial}{\partial x_2} f\right)\right)(c) h_1 h_2 \\
&= \left(\frac{\partial}{\partial x_2}\left(\frac{\partial}{\partial x_1} f\right)\right)(d) h_1 h_2,
\end{aligned} \tag{2.40}$$

where we obtained the last line by symmetry. This shows that

$$\frac{\partial}{\partial x_1}\left(\frac{\partial f}{\partial x_2}\right)(c) = \frac{\partial}{\partial x_2}\left(\frac{\partial f}{\partial x_1}\right)(d). \tag{2.41}$$

As $h \to 0, c \to a$ and $d \to a$, so

$$\frac{\partial}{\partial x_1}\left(\frac{\partial f}{\partial x_2}\right)(a) = \frac{\partial}{\partial x_2}\left(\frac{\partial f}{\partial x_1}\right)(a), \tag{2.42}$$

for any $a \in U$. $\square$

The above argument can be iterated for $f \in \mathcal{C}^3(U)$, $f \in \mathcal{C}^3(4)$, etc.