MA2321—Analysis in Several Variables
School of Mathematics, Trinity College
Michaelmas Term 2017
Lecture 19 (November 20, 2017)

David R. Wilkins

9. Differentiation of Functions of Several Real Variables

9.1. Observations on the Concept of Differentiability

Let us consider the definition of differentiability for functions of a single real variable. Let D be a subset of the set $\mathbb{R}$ of real numbers, and let p be a value in the interior of D. The function f is then said to be differentiable at p, with derivative $f'(p)$, if and only if the limit of the *difference quotient*

$$\frac{f(p+h) - f(p)}{h}$$

as $h \to 0$ exists and is equal to $f'(p)$.

We wish to extend the definition of differentiability to real-valued functions $f\colon D \to \mathbb{R}$ defined on subsets D of m-dimensional Euclidean space $\mathbb{R}^m$. Let $\mathbf{p}$ be a point in the interior of D. Then $\mathbf{p} + \mathbf{h} \in D$ for all m-dimensional vectors $\mathbf{h}$ that are sufficiently close to the zero vector in $\mathbb{R}^m$. Then given any non-zero m-dimensional vector $\mathbf{h}$ for which $\mathbf{p} + \mathbf{h} \in D$, the difference $f(\mathbf{p} + \mathbf{h}) - f(\mathbf{p})$ is a real number. *And, in dimensions greater than two, there is no algebraic operation for dividing real numbers by non-zero vectors that is compatible with the usual "laws" of algebra* such as the Commutative Laws, Associative Laws and Distributive Law satisfied by the operations of addition, subtraction, multiplication and division within the fields of real and complex numbers).

Thus the *derivative of a function of several variables cannot be defined as a limit of difference quotients*.

Next we consider partial derivatives. Let $f\colon D \to \mathbb{R}$ be a function defined on a subset D of $\mathbb{R}^m$. On might be tempted to define differentiability by saying that the function f is differentiable at a point | in the interior of D if and only if all the partial derivatives of f exist at the point $\mathbf{p}$. However *mathematicians do not define differentiability for functions of several variables in this fashion*, and the following example demonstrates why they do not do so.

Example

Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be defined so that

$$f(x,y) = \begin{cases} \dfrac{xy}{(x^2+y^2)^2} & \text{if } (x,y) \neq (0,0); \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

The usual propositions and rules of calculus ensure that

$$\frac{\partial f(x,y)}{\partial x} \quad \text{and} \quad \frac{\partial f(x,y)}{\partial y}$$

are defined for all points (x, y) of $\mathbb{R}^2$ that are distinct from the origin $(0, 0)$. Also

$$\left.\frac{\partial f(x,y)}{\partial x}\right|_{(x,y)=(0,0)} = \lim_{h\to 0} \frac{f(h,0) - f(0,0)}{h} = 0,$$

$$\left.\frac{\partial f(x,y)}{\partial y}\right|_{(x,y)=(0,0)} = \lim_{h\to 0} \frac{f(0,h) - f(0,0)}{h} = 0,$$

because $f(h, 0) = 0$ and $f(0, h) = 0$ for all non-zero real numbers h. Thus the partial derivatives of the function f exist at every point of $\mathbb{R}^2$.

But now let us consider the behaviour of the function f along the line $x = y$. Now

$$f(t,t) = \begin{cases} \dfrac{1}{4t^2} & \text{if } t \neq 0; \\ 0 & \text{if } t = 0. \end{cases}$$

It follows that $f(t,t)$ increases without limit at $t \to 0$, and therefore the function $f\colon \mathbb{R}^2 \to \mathbb{R}$ is not continuous at the origin $(0,0)$.

This example demonstrates that, were mathematicians to take the existence of well-defined partial derivatives as the criterion for differentiability, then many functions would be differentiable that were not continuous.

The following lemma however provides a characterization of differentiability for functions of a single real variable that can be generalized directly so as to apply to functions of several real variables.

Lemma 9.1

Let $f\colon D \to \mathbb{R}$ be a real-valued function defined on some subset D of the set of real numbers. Let s be a real number in the interior of D. The function f is differentiable at s with derivative $f'(s)$ (where $f'(s)$ is some real number) if and only if

$$\lim_{x \to s} \frac{1}{|x - s|} \left(f(x) - f(s) - f'(s)(x - s) \right) = 0.$$

Proof

It follows directly from the definition of the limit of a function that

$$\lim_{x \to s} \frac{f(x) - f(s)}{x - s} = f'(s)$$

if and only if

$$\lim_{x \to s} \left| \frac{f(x) - f(s)}{x - s} - f'(s) \right| = 0.$$

But

$$\left| \frac{f(x) - f(s)}{x - s} - f'(s) \right| = \frac{1}{|x - s|} \left| f(x) - f(s) - f'(s)(x - s) \right|.$$

It follows immediately from this that the function f is differentiable at s with derivative $f'(s)$ if and only if

$$\lim_{x \to s} \frac{1}{|x - s|} \bigl(f(x) - f(s) - f'(s)(x - s) \bigr) = 0. \qquad \blacksquare$$

Now let us observe that, for any real number c, the map $h \mapsto ch$ defines a linear transformation from $\mathbb{R}$ to $\mathbb{R}$. Conversely, every linear transformation from $\mathbb{R}$ to $\mathbb{R}$ is of the form $h \mapsto ch$ for some $c \in \mathbb{R}$. Because of this, we may regard the derivative $f'(s)$ of f at s as representing a linear transformation $h \mapsto f'(s)h$, characterized by the property that the map

$$x \mapsto f(s) + f'(s)(x - s)$$

provides a 'good' approximation to f around s in the sense that

$$\lim_{x \to s} \frac{e(x)}{|x - s|} = 0,$$

where

$$e(x) = f(x) - f(s) - f'(s)(x - s)$$

(i.e., $e(x)$ measures the difference between $f(x)$ and the value $f(s) + f'(s)(x - s)$ of the approximation at x, and thus provides a measure of the error of this approximation).

We shall generalize the notion of differentiability to functions φ from $\mathbb{R}^m$ to $\mathbb{R}^n$ by defining the derivative $(D\varphi)_{\mathbf{p}}$ of φ at $\mathbf{p}$ to be a linear transformation from $\mathbb{R}^m$ to $\mathbb{R}^n$ characterized by the property that the map

$$\mathbf{x} \mapsto \varphi(\mathbf{p}) + (D\varphi)_{\mathbf{p}}\,(\mathbf{x} - \mathbf{p})$$

provides a 'good' approximation to φ around $\mathbf{p}$.

9.2. Derivatives of Functions of Several Variables

Definition

Let X be an open subset of $\mathbb{R}^m$ and let $\varphi\colon X \to \mathbb{R}^n$ be a map from X into $\mathbb{R}^n$. Let $\mathbf{p}$ be a point of X. The function φ is said to be *differentiable* at $\mathbf{p}$, with *derivative* $(D\varphi)_{\mathbf{p}}\colon \mathbb{R}^m \to \mathbb{R}^n$, where $(D\varphi)_{\mathbf{p}}$ is a linear transformation from $\mathbb{R}^m$ to $\mathbb{R}^n$, if and only if

$$\lim_{\mathbf{x}\to\mathbf{p}} \frac{1}{|\mathbf{x}-\mathbf{p}|} \left(\varphi(\mathbf{x}) - \varphi(\mathbf{p}) - (D\varphi)_{\mathbf{p}}(\mathbf{x}-\mathbf{p})\right) = \mathbf{0}.$$

The derivative of a map $\varphi\colon X \to \mathbb{R}^n$ defined on a open subset X of $\mathbb{R}^m$ at a point $\mathbf{p}$ of X is usually denoted either by $(D\varphi)_{\mathbf{p}}$ or else by $\varphi'(\mathbf{p})$.

The derivative $(D\varphi)_{\mathbf{p}}$ of φ at $\mathbf{p}$ is sometimes referred to as the *total derivative* of φ at $\mathbf{p}$. If φ is differentiable at every point of X then we say that φ is differentiable on X.

Lemma 9.2

Let $T\colon \mathbb{R}^m \to \mathbb{R}^n$ be a linear transformation from $\mathbb{R}^m$ into $\mathbb{R}^n$. Then T is differentiable at each point $\mathbf{p}$ of $\mathbb{R}^m$, and $(DT)_{\mathbf{p}} = T$.

Proof

This follows immediately from the identity
$T\mathbf{x} - T\mathbf{p} - T(\mathbf{x} - \mathbf{p}) = \mathbf{0}$. ∎

Lemma 9.3

Let $\varphi\colon X \to \mathbb{R}^n$ be a function, let $T\colon \mathbb{R}^m \to \mathbb{R}^n$ be a linear transformation, and let $\Omega\colon X \to \mathbb{R}^n$ be defined so that

$$\Omega(\mathbf{x}) = \begin{cases} \dfrac{1}{|\mathbf{x}-\mathbf{p}|}\left(\varphi(\mathbf{x}) - \varphi(\mathbf{p}) - T(\mathbf{x}-\mathbf{p})\right) & \text{if } \mathbf{x} \neq \mathbf{p}; \\ \mathbf{0} & \text{if } \mathbf{x} = \mathbf{p}. \end{cases}$$

Then $\varphi\colon X \to \mathbb{R}^n$ is differentiable at $\mathbf{p}$ with derivative $T\colon \mathbb{R}^m \to \mathbb{R}^n$ if and only if $\lim_{\mathbf{x}\to\mathbf{p}} \Omega(\mathbf{x}) = \mathbf{0} = \Omega(\mathbf{p})$. Thus the function $\varphi\colon X \to \mathbb{R}^n$ is differentiable at $\mathbf{p}$, with derivative T, if and only if the associated function $\Omega\colon X \to \mathbb{R}^n$ is continuous at $\mathbf{p}$

Proof

It follows from the definition of differentiability that the function φ is differentiable, with derivative $T\colon \mathbb{R}^m \to \mathbb{R}^n$, if and only if $\lim_{\mathbf{x}\to\mathbf{p}} \Omega(\mathbf{x}) = \mathbf{0}$. But $\Omega(\mathbf{p}) = \mathbf{0}$. It follows that $\lim_{\mathbf{x}\to\mathbf{p}} \Omega(\mathbf{x}) = \mathbf{0}$ if and only if the function Ω is continuous at $\mathbf{p}$ (see Proposition 4.17). The result follows. ∎

Example
Let $\varphi\colon \mathbb{R}^2 \to \mathbb{R}^2$ be defined so that

$$\varphi\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} x^2 - y^2 \\ 2xy \end{pmatrix}$$

for all real numbers x and y. Let p and q be fixed real numbers. Then

$$
\begin{aligned}
\varphi\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) - \varphi\left(\begin{pmatrix} p \\ q \end{pmatrix}\right)
&= \begin{pmatrix} x^2 - y^2 \\ 2xy \end{pmatrix} - \begin{pmatrix} p^2 - q^2 \\ 2pq \end{pmatrix} \\
&= \begin{pmatrix} (x+p)(x-p) - (y+q)(y-q) \\ 2q(x-p) + 2p(y-q) + 2(x-p)(y-q) \end{pmatrix} \\
&= \begin{pmatrix} 2p(x-p) - 2q(y-q) + (x-p)^2 - (y-q)^2 \\ 2q(x-p) + 2p(y-q) + 2(x-p)(y-q) \end{pmatrix} \\
&= \begin{pmatrix} 2p & -2q \\ 2q & 2p \end{pmatrix} \begin{pmatrix} x-p \\ y-q \end{pmatrix} + \begin{pmatrix} (x-p)^2 - (y-q)^2 \\ 2(x-p)(y-q) \end{pmatrix}.
\end{aligned}
$$

Now, given $(x, y) \in \mathbb{R}^2$, let $r = \sqrt{(x-p)^2 + (y-q)^2}$. Then $|x - p| < r$ and $|y - q| < r$, and therefore

$$|(x-p)^2 - (y-q)^2| \leq |x-p|^2 + |y-q|^2 < 2r^2$$

and $2(x-p)(y-q) < 2r^2$, and thus

$$\frac{(x-p)^2 - (y-q)^2}{\sqrt{(x-p)^2 + (y-q)^2}} < 2r \quad \text{and} \quad \frac{2(x-p)(y-q)}{\sqrt{(x-p)^2 + (y-q)^2}} < 2r.$$

Thus, given any positive real number ε, let $\delta = \frac{1}{2}\varepsilon$. Then

$$\left| \frac{(x-p)^2 - (y-q)^2}{\sqrt{(x-p)^2 + (y-q)^2}} \right| < \varepsilon \quad \text{and} \quad \left| \frac{2(x-p)(y-q)}{\sqrt{(x-p)^2 + (y-q)^2}} \right| < \varepsilon$$

whenever $0 < |(x, y) - (p, q)| < \delta$. It follows therefore that

$$\lim_{(x,y)\to(0,0)} \frac{1}{\sqrt{(x-p)^2+(y-q)^2}} \left(\begin{array}{c} (x-p)^2-(y-q)^2 \\ 2(x-p)(y-q) \end{array} \right) = \left(\begin{array}{c} 0 \\ 0 \end{array} \right).$$

Thus the function $\varphi\colon \mathbb{R}^2 \to \mathbb{R}^2$ is differentiable at (p,q), and the derivative of this function at (p,q) is the linear transformation represented by the matrix

$$\left(\begin{array}{cc} 2p & -2q \\ 2q & 2p \end{array} \right).$$