Differentiability for Multivariable Functions

Louis A. Talman, Ph.D.
Department of Mathematical & Computer Sciences
Metropolitan State College of Denver

September 12, 2007

In order to understand the notion of differentiability as it is usually defined for functions of several variables, it is helpful to develop first a deeper understanding than is often required in the elementary calculus sequence of just what differentiability means for functions of a single variable. Ordinarily, in a first calculus course, and in the two or three ensuing courses, we understand differentiability of a function f of a single variable at the point a to mean simply that

$$f'(a) = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h} \tag{1}$$

exists. A consequence of the existence of this limit is the existence of a function, φ, with the properties

1. $\lim_{h \to 0} \varphi(h) = 0$, and
2. $f(a+h) = f(a) + f'(a)h + \varphi(h) \cdot h$ for all non-zero values of h that are sufficiently close to 0.

The existence of such a φ is sometimes called *the fundamental increment lemma,* and we prove it simply by putting

$$\varphi(h) = \frac{f(a+h) - f(a)}{h} - f'(a) \tag{2}$$

and noting that the function so defined has the necessary properties. This result proves its usefulness when we need to prove the Chain Rule—among other things.

What we rarely point out in the early calculus courses is that the conclusion of the fundamental increment lemma, properly restated, essentially *characterizes* differentiability for functions of a single variable. For suppose that there are a number m and a function φ defined on some deleted neighborhood of 0 such that

1. $\lim_{h\to 0} \varphi(h) = 0$, and
2. $f(a+h) = f(a) + mh + \varphi(h)\cdot h$ for all non-zero values of h that are sufficiently close to 0.

Then, when $h \neq 0$, we have

$$\begin{aligned} \frac{f(a+h)-f(a)}{h} &= \frac{mh+\varphi(h)\cdot h}{h} && (3) \\ &= m + \varphi(h), && (4) \end{aligned}$$

and it is clear that $\lim_{h\to 0}[m + \varphi(h)] = m$. Thus, the existence of a number m and the function φ having together the properties described guarantee that $f'(a)$ exists and is m.

Now let us rephrase this equivalent description of differentiability for functions of a single variable in a way that we allows us to interpret it for functions of several variables. The key to doing this is the notion of a *linear function.*

Definition 1: *A function* $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^m$ *is called a* linear function *if, for every pair of real numbers* α *and* β *and every pair of* n*-vectors* $\mathbf{u}$ *and* $\mathbf{v}$ *we have*

$$\mathbf{F}(\alpha\mathbf{u} + \beta\mathbf{v}) = \alpha\mathbf{F}(\mathbf{u}) + \beta\mathbf{F}(\mathbf{v}). \quad (5)$$

Note 1: We have an unfortunate habit of calling functions of a single variable "linear" if they have the form $f(x) = mx + b$. We shouldn't do that, because these functions don't meet the requirements of the definition above. We ought to call these functions "affine" functions.

Note 2: You shouldn't have much trouble convincing yourself that the only linear functions, according to the definition above, from $\mathbb{R}$ to $\mathbb{R}$ are those of the form $M(x) = mx$, where m is a constant.

Using the language of linear functions, we can rephrase the differentiability property for functions of a single variable in yet another way: A function $f : \mathbb{R} \to \mathbb{R}$ is differentiable iff there is a linear function $M : \mathbb{R} \to \mathbb{R}$ and a function φ defined in some deleted neighborhood of 0 such that

1. $\lim_{h\to 0} \varphi(h) = 0$, and
2. $f(a+h) = f(a) + M(h) + \varphi(h)\cdot h$ for all non-zero values of h that are sufficiently close to 0.

It is easy to extend this formulation of differentiability to functions $f : D \to \mathbb{R}$*, where* $D \subseteq \mathbb{R}^n$*.*

Definition 2: *A function* $f : D \to \mathbb{R}$*, where* $D \subseteq \mathbb{R}^n$*, is said to be* differentiable at the point $\mathbf{a} \in D$ *if there is a linear function* $M : \mathbb{R}^n \to \mathbb{R}$ *and a function* Φ*, carrying some deleted neighborhood of the origin in* $\mathbb{R}^n$ *to* $\mathbb{R}^n$*, such that*

1. $\lim_{\mathbf{h}\to\mathbf{0}} \varphi(\mathbf{h}) = \mathbf{0}$*, and*
2. $f(\mathbf{a}+\mathbf{h}) = f(\mathbf{a}) + M(\mathbf{h}) + \Phi(\mathbf{h}) \cdot \mathbf{h}$ *for all non-zero vectors* $\mathbf{h}$ *that are sufficiently close to* $\mathbf{0} \in \mathbb{R}^n$*.*

When all of these things are so, the linear function M *is called* the derivative of f at a*. It is sometimes written* Df_a*.*

Definition 2 says, essentially, that we call a function "differentiable" at a point precisely when its increments there are well-approximated by a linear function acting on the corresponding (vector) increments of the independent variable.

Now the linear functions acting from $\mathbb{R}^2$ to $\mathbb{R}$ all have the form $(h, k) \mapsto ph+qk$, where p and q are constants, so for functions f acting from some set $D \subseteq \mathbb{R}^2$ into $\mathbb{R}$ the requirement of Definition 2 for differentiability at a point (x_0, y_0) translates to the existence of two numbers, p and q, and two functions, φ and ψ, both of two variables and defined in some deleted neighborhood of (x_0, y_0), such that

1. $\lim_{(h,k)\to(0,0)} \varphi(h, k) = 0$,
2. $\lim_{(h,k)\to(0,0)} \psi(h, k) = 0$, and
3. $f(x_0 + h, y_0 + k) = f(x_0, y_0) + ph + qk + \varphi(h, k)h + \psi(h, k)k$ for all non-zero (h, k) that are sufficiently close to $(0, 0)$.

The graph, in three dimensions, of the function $(h, k) \mapsto f(x_0, y_0)+ph+qk$ is, of course, the graph of a certain plane which passes through the point $\big(x_0, y_0, f(x_0, y_0)\big)$. So for such functions, the differentiability requirement is that the surface $z = f(x, y)$ be well-approximated near the point (x_0, y_0) by a certain plane—which we call the tangent plane.

It is not difficult to show that if the numbers p and q and the functions φ and ψ of the previous paragraph exist, then the partial derivatives $f_1(x_0, y_0)$ and $f_2(x_0, y_0)$ exist and are, respectively, p and q. *If a function* f *of two variables is differentiable at* (x_0, y_0)*, then it possesses both of its partial derivatives,* and, indeed, possesses all of its directional derivatives, at (x_0, y_0). We can also show, though it requires substantially more sophistication and work to do so[1], that when the partial derivatives of f are both continuous in a neighborhood of (x_0, y_0), the function f must be differentiable at (x_0, y_0).

[1]We must first prove a mean value theorem for functions of two variables. Unlike the MVT for functions of a single variable, this MVT requires that the (partial) derivatives be continuous in the region where we are working.

But here the parallels with the single-variable theory end. In particular, the mere existence of the partial derivatives $f_1(x_0, y_0)$ and $f_2(x_0, y_0)$ *does not imply* that f is differentiable at (x_0, y_0), and even the existence of *all* of the directional derivatives[2] isn't enough. This can be seen by considering the example, to which we shall turn directly, of the function g given by

$$g(x,y) = \begin{cases} \dfrac{x^3}{x^2+y^2}, & \text{when } x^2+y^2 \neq 0, \\ \\ 0, & \text{when } x^2+y^2 = 0. \end{cases} \tag{6}$$

The underlying difficulty is that partial (or directional) derivatives depend only upon the values of the function along a single line, but functions of two variables can twist and bend enough to defy any analysis that depends only on information gathered along straight lines—even if we consider every straight line that passes through the particular point where we are carrying out our analysis. The extra room in which functions can wiggle and misbehave makes life with multivariable calculus more interesting than life with single variable calculus was.

When g is as given in the previous paragraph, and as long as $x^2 + y^2 \neq 0$, we have, by the usual differentiation rules,

$$g_1(x,y) = \frac{x^4 + 3x^2y^2}{(x^2+y^2)^2}, \text{ and} \tag{7}$$

$$g_2(x,y) = -\frac{2x^3y}{(x^2+y^2)^2}. \tag{8}$$

The algebraic expression for $g(x, y)$ is meaningless when $x^2 + y^2 = 0$, so we must have recourse to limits of difference quotients at the origin, where $g(0, 0) = 0$ by definition. We have

$$g_1(0,0) = \lim_{h\to 0} \frac{g(0+h,0) - g(0,0)}{h} \tag{9}$$

$$= \lim_{h\to 0} \frac{1}{h}\left(\frac{h^3}{h^2+0^2} - 0\right) = \lim_{h\to 0} \frac{h^3}{h^3} = \lim_{h\to 0} 1 = 1, \tag{10}$$

while

$$g_2(0,0) = \lim_{k\to 0} \frac{g(0,0+k) - g(0,0)}{k} \tag{11}$$

$$= \lim_{k\to 0} \frac{1}{k}\left(\frac{0^3}{0^2+k^2} - 0\right) = \lim_{k\to 0} \left(\frac{1}{k}\cdot 0\right) = \lim_{k\to 0} 0 = 0. \tag{12}$$

[2] The partial derivatives are special cases of the directional derivatives.

The partial derivatives g_1 and g_2 thus exist everywhere.

However, $\lim_{(x,y)\to(0,0)} g_1(x,y)$ and $\lim_{(x,y)\to(0,0)} g_2(x,y)$ do not exist, so that neither of the partial derivatives of g is continuous at $(0,0)$. In order to see this, consider what happens in each case if we force (x,y) to approach $(0,0)$ along the straight line $y = mx$. We then have

$$\lim_{(x,y)\to(0,0)} g_1(x,y) = \lim_{x\to 0} g_1(x,mx) \tag{13}$$

$$= \lim_{x\to 0} \frac{x^4 + 3m^2x^4}{(x^2 + m^2x^2)^2} \tag{14}$$

$$= \lim_{x\to 0} \frac{1+3m^2}{(1+m^2)^2} = \frac{1+3m^2}{(1+m^2)^2} \tag{15}$$

and

$$\lim_{(x,y)\to(0,0)} g_2(x,y) = \lim_{x\to 0} g_2(x,mx) \tag{16}$$

$$= -\lim_{x\to 0} \frac{2mx^4}{(x^2 + m^2x^2)^2} \tag{17}$$

$$= -\lim_{x\to 0} \frac{2m}{(1+m^2)^2} = -\frac{2m}{(1+m^2)^2}. \tag{18}$$

Thus, in both cases, the limits as we approach the origin along different lines have different values. Consequently, neither of the limits as $(x,y) \to (0,0)$ exists. This means that neither partial derivative is continuous at the origin.

Is g differentiable at the origin? To answer this question we must determine whether or not there are a pair of numbers p and q and a pair of functions φ and ψ, all with the properties required by our translation of Definition 2 for functions of two variables. From our earlier discussion, we know that if such numbers and functions exist, we will have to have $p = g_1(0,0) = 1$ and $q = g_2(0,0) = 0$. But

$$g(0+h, 0+k) - g(0,0) - 1\cdot h - 0\cdot k = \frac{h^3}{h^2+k^2} - h - 0\cdot k \tag{19}$$

$$= \left(\frac{h^2}{h^2+k^2} - 1\right)\cdot h + 0\cdot k \tag{20}$$

$$= -\frac{k^2}{h^2+k^2}\cdot h + 0\cdot k. \tag{21}$$

In order for g to be differentiable at the origin, we must therefore have

$$\varphi(h,k) = -\frac{k^2}{h^2+k^2}, \text{ and} \tag{22}$$

$$\psi(h,k) = 0. \tag{23}$$

The function ψ causes no trouble. But φ is quite another matter, because—again considering what happens as we approach the origin along straight lines, this time of the form $k = mh$—we have

$$\lim_{(h,k)\to(0,0)} \varphi(h,k) = \lim_{h\to 0} \varphi(h, mh) = -\lim_{h\to 0} \frac{m^2h^2}{h^2 + m^2h^2} \tag{24}$$

$$= -\lim_{h\to 0} \frac{m^2}{1+m^2} = -\frac{m^2}{1+m^2}, \tag{25}$$

which once again gives different results along different lines so that the limit as $(h,k) \to (0,0)$ of $\varphi(h,k)$ doesn't exist—and consequently can't be zero as is required for differentiability. We are forced to conclude that the function g is not differentiable at the origin, or, more intuitively, that no plane through the origin gives a suitable approximation to the surface $z = g(x,y)$ at the point $(0,0,0)$, which lies on the surface.

Finally, let us remark that although continuity of the first order partial derivatives of g in some neighborhood of the point (x_0, y_0) guarantees that g is differentiable at (x_0, y_0), it is possible for a function to be differentiable at (x_0, y_0) when its partial derivatives are not continuous there. In order to see that this is so, let g now denote the function given by

$$g(x,y) = \begin{cases} (x^2+y^2)\sin\left(\dfrac{1}{\sqrt{x^2+y^2}}\right), & \text{when } x^2+y^2 \neq 0, \\ \\ 0, & \text{when } x^2+y^2 = 0. \end{cases} \tag{26}$$

We show first that g is differentiable by taking $p = q = 0$ and putting

$$\varphi(h,k) = h\sin\left(\frac{1}{\sqrt{h^2+k^2}}\right) \text{ and} \tag{27}$$

$$\psi(h,k) = k\sin\left(\frac{1}{\sqrt{h^2+k^2}}\right). \tag{28}$$

We note that

$$\left| h\sin\left(\frac{1}{\sqrt{h^2+k^2}}\right)\right| \leq |h| \to 0 \tag{29}$$

as $(h,k) \to (0,0)$, so that $\lim_{(h,k)\to(0,0)} \varphi(h,k) = 0$. Similarly, $\lim_{(h,k)\to(0,0)} \psi(h,k) = 0$.

Then we note that

$$g(0+h, 0+k) = (h^2+k^2)\sin\left(\frac{1}{\sqrt{h^2+k^2}}\right) \tag{30}$$

$$= 0 + 0\cdot h + 0\cdot k + \left[h\sin\left(\frac{1}{\sqrt{h^2+k^2}}\right)\right]h + \left[k\sin\left(\frac{1}{\sqrt{h^2+k^2}}\right)\right]k \tag{31}$$

$$= 0 + 0\cdot h + 0\cdot k + \varphi(h,k)h + \psi(h,k)k \tag{32}$$

$$= g(0,0) + 0\cdot h + 0\cdot k + \varphi(h,k)h + \psi(h,k)k. \tag{33}$$

We have written $g(0+h, 0+k)$ in the required form, and we must now conclude that g is differentiable at $(0,0)$. We also conclude, from what we said earlier, that $g_1(0,0) = p = 0$ and that $g_2(0,0) = q = 0$.

However, when $x^2+y^2 \neq 0$, we have, by standard differentiation techniques,

$$g_1(x,y) = -\frac{x\cos[(x^2+y^2)^{-1/2}]}{\sqrt{x^2+y^2}} + 2x\sin[(x^2+y^2)^{-1/2}]. \tag{34}$$

It is easily seen that $\lim_{(x,y)\to(0,0)} g_1(x,y)$ does not exist: We have already seen that

$$\lim_{(x,y)\to(0,0)} x\sin\left[\frac{1}{\sqrt{x^2+y^2}}\right] = 0, \tag{35}$$

so the second piece of this derivative is well-behaved. But, calling the $y = mx$ trick into play again, we would have to have

$$\lim_{(x,y)\to(0,0)} \frac{x\cos[(x^2+y^2)^{-1/2}]}{\sqrt{x^2+y^2}} = \lim_{x\to 0} \frac{x\cos[(x^2+m^2x^2)^{-1/2}]}{\sqrt{x^2+m^2x^2}} \tag{36}$$

$$= \lim_{x\to 0}\left[\frac{x}{|x|}\cdot\frac{\cos[(x^2+m^2x^2)^{-1/2}]}{\sqrt{1+m^2}}\right], \tag{37}$$

which does not exist owing to the fact that the fraction involving the cosine oscillates infinitely many times, from $1/\sqrt{1+m^2}$ to $-1/\sqrt{1+m^2}$ and back again in every open interval which has 0 as an endpoint. Similarly for $\lim_{(x,y)\to(0,0)} g_2(x,y)$. Thus, the partial derivatives of g fail of continuity at the origin.