DIFFERENTIABILITY FOR FUNCTIONS OF SEVERAL VARIABLES

We begin by reviewing the concept of differentiation for functions of one variable.

Definition 1. *Let $f : D \subset \mathbb{R} \to \mathbb{R}$ and let a be an interior point of D. Then f is differentiable at a means*

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

or equivalently

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

exists. The number $f'(a)$ is called the derivative of f at a.

Geometrically the derivative of a function at a is interpreted as the slope of the line tangent to the graph of f at the point $(a, f(a))$. Not every function is differentiable at every number in its domain even if that function is continuous. For example $f(x) = |x|$ is not differentiable at 0 but f is continuous at 0. However we do have the following theorem.

Theorem 1. *If f is differentiable at a, then f is continuous at a.*

Extending the definition of differentiability in its present form to functions of several variables is not possible because the definition involves division and dividing by a vector or by a point in n dimensional space is not possible. To carry out the extension an equivalent definition is developed that involves division by a distance. The limit statement can be rewritten as

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a} - f'(a) = 0 \text{ or } \lim_{x \to a} \frac{f(x) - f(a) - (x-a)f'(a)}{x - a} = 0.$$

One final modification is still necessary.

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

So the following definition is equivalent to the original one.

Definition 2. *Let $f : D \subset \mathbb{R} \to \mathbb{R}$ and let a be an interior point of D. Then f is differentiable at a means there is a number, $f'(a)$, such that*

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

One way to interpret this expression is that $f(x) - f(a) - (x-a)f'(a)$ tends to 0 faster than $|x-a|$ does and consequently $f(x)$ is approximately equal to $f(a) + (x-a)f'(a)$. The equation $y = f(a) + (x-a)f'(a)$ is the equation of the line tangent to the graph of f at the point $(a, f(a))$. So $f(x)$ is approximated very well by its tangent line. This observation is the bases for linear approximation.

Using this form of the definition as a model it is possible to construct a definition of differentiability for functions of several variables. What goes in the denominator is fairly easy to see; namely, $|P - P_0|$. Similarly the first two term in the numerator would become $f(P) - f(P_0)$. But what should replace the term $(x-a)f'(a)$? First we note that it must be a number. One of the factors will be $(P_0 - P)$ or better yet $(\overrightarrow{P_0P})$ — a vector. Consequently the other must also be a vector and the product will be the dot product. With these observation the definition of differentiability for functions of several variable is as follows.

Definition 3. *Let $f : D \subset \mathbb{R}^n \to \mathbb{R}$ and let P_0 be an interior point of D. Then f is differentiable at P_0 means there is a vector, which will be denoted by $f'(P_0)$ for now, such that*

$$\lim_{P \to P_0} \frac{f(P) - f(P_0) - (\overrightarrow{P_0P}) \cdot f'(P_0)}{|P - P_0|} = 0.$$

For functions of two variables the definition becomes the following.

Definition 4. *Let $f : D \subset \mathbb{R}^2 \to \mathbb{R}$ and let (x_0, y_0) be an interior point of D. Then f is differentiable at (x_0, y_0) means there are two numbers, $f_1(x_0, y_0)$ and $f_2(x_0, y_0)$ such that*

$$\lim_{(x,y)\to(x_0,y_0)} \frac{f(x,y) - f(x_0,y_0) - (x - x_0)f_1(x_0,y_0) - (y - y_0)f_2(x_0,y_0)}{\sqrt{(x - x_0)^2 + (y - y_0)^2}} = 0. \tag{1}$$

The vector $f_1(x_0, y_0)\mathbf{i} + f_2(x_0, y_0)\mathbf{j}$ or the pair $(f_1(x_0, y_0), f_2(x_0, y_0))$ is called the derivative of f at the point (x_0, y_0).

Interpret this definition as requiring that the graph of f have a tangent plane at the point $(x_0, y_0, f(x_0, y_0))$. In fact it is easy to get an equation for this tangent plane. It is

$$z = f(x_0, y_0) + f_1(x_0, y_0)(x - x_0) + f_2(x_0, y_0)(y - y_0).$$

A vector normal to this plane is $f_1(x_0, y_0)\mathbf{i} + f_2(x_0, y_0)\mathbf{j} - \mathbf{k}$. The two numbers, $f_1(x_0, y_0)$ and $f_2(x_0, y_0)$, are computed using techniques learned for computing derivatives of functions of one variable. To find $f_1(x_0, y_0)$ let $y = y_0$ in equation (1) of Definition 4. We get

$$\lim_{x \to x_0} \frac{f(x, y_0) - f(x_0, y_0) - (x - x_0)f_1(x_0, y_0)}{|x - x_0|} = 0.$$

Comparing this statement to Definition 2 we see that $f_1(x_0, y_0)$ is the derivative of the function $h(x) = f(x, y_0)$ at x_0. For example suppose $f(x, y) = x^2y + xy^3$. Then $h(x) = x^2y_0 + xy_0^3$. Differentiating h with respect to x; that is, treating y_0 as a constant, we get that $f_1(x_0, y_0) = 2x_0y_0 + y_0^3$ or more generally for each (x, y) we have $f_1(x, y) = 2xy + y^3$. Notice that this equation is obtained by differentiating the formula for $f(x, y)$ with respect to x treating y as if it were a constant. In a similar fashion $f_2(x, y)$ is obtained by differentiating the formula for $f(x, y)$ with respect to y treating x as a constant. So for the preceding example we get $f_2(x, y) = x^2 + 3xy^2$.

We call $f_1(x, y)$ the first order partial derivative of f with respect to x (or with respect to the first variable) and $f_2(x, y)$ the first order partial derivative of f with respect to y (or with respect to the second variable). As the word, "first" indicates, there are "second", "third" etc. order partial derivatives as well. We will discuss them later. When it is clear that we are dealing with first order partial derivatives the word "first" is often omitted. The following notation is also used to denote partial derivatives.

$$f_1(x, y) = \frac{\partial f}{\partial x}(x, y) = \frac{\partial}{\partial x} f(x, y) = f_x(x, y).$$

In the last expression the same symbol x is use for two different purposes. First, as a subscript where it denotes the variable of differentiation and second, as the first coordinate of a point in $\mathbb{R}^2$. Strictly speaking such a dual use of one symbol is improper, but this abuse is so common as to be acceptable. As one would expect there is analogous notation for $f_2(x, y)$.

The situation for functions of more than two variables is analogous. In the general case, the derivative is a vector in n space and it is computed by computing all of the first order partial derivatives.

As in the case of functions of one variable, differentiability implies continuity.

Theorem 2. *Let $f : D \subset \mathbb{R}^n \to \mathbb{R}$ and let P_0 be an interior point of D. Suppose f is differentiable at P_0. Then f is continuous at P_0.*

Proof: First write

$$\begin{aligned} f(P) - f(P_0) &= f(P) - f(P_0) - (P - P_0) \cdot \overrightarrow{P_0P} + (P - P_0) \cdot \overrightarrow{P_0P} \\ &= \frac{f(P) - f(P_0) - (P - P_0) \cdot \overrightarrow{P_0P}}{|P - P_0|} |P - P_0| + (P - P_0) \cdot \overrightarrow{P_0P} \end{aligned}$$

Since both terms on the right hand side have limit 0 as $P \to P_0$,

$$\lim_{P \to P_0} f(P) - f(P_0) = 0; \text{ that is, } \lim_{P \to P_0} f(P) = f(P_0).$$

The converse of the preceding theorem is not true since the converse of the analogous theorem for functions of one variable is not true.

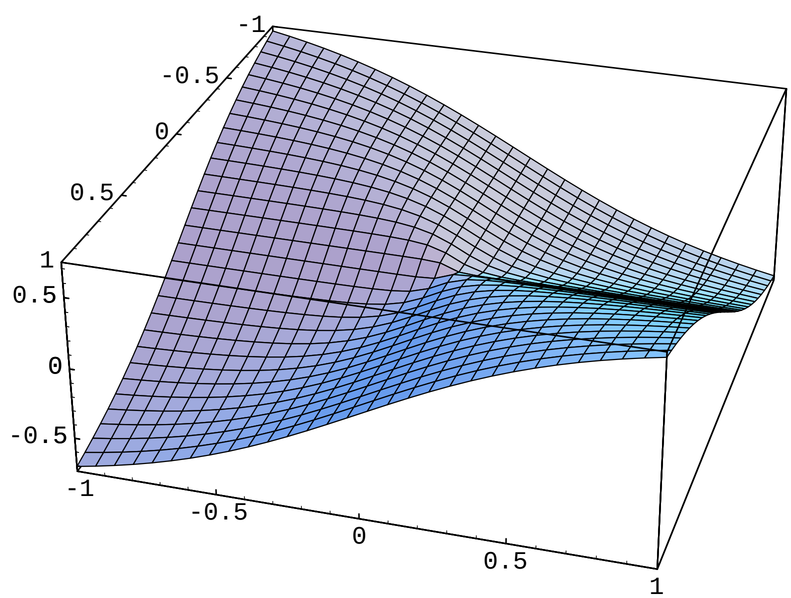
The analogy between differentiation for functions of one variable and for functions of several variable is not a total analogy. For functions of one variable if the derivative, $f'(x)$, can be computed, then f is differentiable at x. The corresponding assertion for functions of two variables is false which stands to reason after considering for a moment what it takes to compute the derivative, $(f_1(x,y), f_2(x,y))$, of a function of two variable. To find $f_1(x_0, y_0)$ one need only know the values of the function, f, at points on the line $y = y_0$ and to find $f_2(x_0, y_0)$ one need only know the values of f at points on the line $x = x_0$. Consequently, the values of f at points not on these two lines play no role in determining the derivative of f. However these values certainly are taken into account when determining whether or not f is differentiable at (x_0, y_0); that is, if the graph of f has a tangent plane at the point (x_0, y_0). For example let

$$f(x) = \begin{cases} 0 & \text{if } x = 0 \text{ or } y = 0 \\ 1 & \text{otherwise} . \end{cases}$$

Since f is 0 on the two coordinate axes, $f_1(0,0) = 0 = f_2(0,0)$ but f is not continuous at $(0,0)$ and by the preceding theorem, f can't be differentiable at $(0,0)$. You might suspect that if f is continuous at (x_0, y_0) and the first order partial derivatives exist there, then f is differentiable at (x_0, y_0) but that conjecture is false as the following example shows. Let

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

The graph of f is pictured below.



Again since f is 0 on the two coordinate axes, $f_1(0,0) = 0 = f_2(0,0)$. So if f were differentiable at $(0,0)$, we would have that

$$\lim_{(x,y)\to(x_0,y_0)} \frac{f(x,y)}{\sqrt{x^2+y^2}} = 0; \text{ that is, } \lim_{(x,y)\to(x_0,y_0)} \frac{xy}{x^2+y^2} = 0.$$

But if the limit is computed along the path $y = x$, we get $\lim_{x\to 0} \frac{x^2}{2x^2} = \frac{1}{2}$.

The natural question to ask then is under what conditions can we conclude that f is differentiable at (x,y). The answer is contained in the following theorem.

Theorem 3. *Let $f : D \subset \mathbb{R}^n \to \mathbb{R}$ and let P_0 be an interior point of D. Suppose all n of the first order partial derivatives of f exist in a ball about P_0 and are continuous at P_0. Then f is differentiable at P_0.*

For example let $f(x, y) = \sqrt{y^2 - x^2} = (y^2 - x^2)^{1/2}$. Then $f_1(x, y) = -x(y^2 - x^2)^{-1/2}$ and $f_2(x, y) = y(y^2 - x^2)^{-1/2}$. These two functions are continuous in the region consisting of that part of $\mathbb{R}^2$ above the graph of $y = |x|$ together with that part of $\mathbb{R}^2$ below the graph of $y = -|x|$. According to the theorem, f is differentiable on this region.