

Math 410/610: (2/18/91) The Total Differential of a Function

You may not realize it, but we have not yet discussed the concept of differentiability for a function of more than one variable! How can I say this when we have already discussed partial derivatives, directional derivatives and even finding the tangent plane to parametrized surfaces and to graphs of functions of several variables? The point is a subtle one, but to put it simply, the existence of partial derivatives only says the function is "partially differentiable", whereas to say a function is (fully) differentiable means that the (total) differential exists.

The difference is slightly confusing because if the total differential does exist, then it is completely determined by the partial derivatives. Think of an analogous question, that of determining the equation for a surface S which passes through the origin in \mathbb{R}^3 . Suppose we know that S contains both the x -axis and the y -axis and suppose we know that S is a plane. Then it follows that S must be the (x,y) plane. But suppose we only know that S contains the x -axis and the y -axis, but we do not know that S is a plane. Does S have to be the (x,y) plane? Or is there some other surface which is not a plane but still contains the x and y -axes? In fact there are very many other such surfaces as you will realize if you think about it for a while. [For example the surface in \mathbb{R}^3 with equation $z = xy$ contains all points with both $z=0$ and $y=0$, hence contains the x -axis, and also all points with $z=0$ and $x=0$, hence contains the y -axis. It is not a plane however, since its equation is not linear, and in particular it is not the (x,y) plane since it does not contain the point $(0,0,1)$.]

The point is that if you already know a surface is a plane then just knowing two lines in it tells you completely which plane it is, and allows you to write an equation for that plane. On the other hand if you don't really know whether it is a plane or not, then just knowing that there are two lines in it does not really help you know what the surface looks like in other directions. This is a very good analogy of the difference between a function which has partial derivatives and a function which is (fully) differentiable. I.e. if the function is differentiable then just knowing the partial derivatives tells you what the (total) derivative is, but a function can have partial derivatives and not be differentiable at all.

The connection is this: for a function $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ to be "differentiable at p " will turn out to mean that the graph has a tangent plane at the point $(p, f(p))$. This implies that

2

- i) every curve lying in the graph of f and passing through the point $(p, f(p))$ has a velocity vector, and
- (ii) the set of all such velocity vectors lie in a common plane.

Thus at least two things that can go wrong, either some curves in the graph through $(p, f(p))$ have velocity vectors and some do not, or all such curves have velocity vectors but those vectors do not all lie in a common plane. For f to have partial derivatives at p simply means that at least two curves through $(p, f(p))$ have velocity vectors, namely the two curves in $\text{graph}(f)$ which move in the direction of the x axis and the y axis. This does not at all guarantee that curves in other directions will have velocity vectors, much less that all of them will lie in a common plane.

For example, if we define $f(x, y) = 0$ when either x or y is 0, and $f(x, y) = 1$ when neither x nor y is 0, then this f has both partial derivatives at $(0, 0)$, namely $\partial f / \partial x = 0$, and $\partial f / \partial y = 0$, but f is not even continuous at $(0, 0)$, and no curve through the origin in any direction except along the two axes has a velocity vector there. The graph looks like the set you would get if you tried to lift the (x, y) plane up one unit but somebody had glued the x and y axes down, so they stuck where they were while the rest of the plane ripped loose and came up a distance of one unit. In particular f is not differentiable at $p = (0, 0)$. This function does not satisfy either condition i) or condition ii) above.

For a function f to have a directional derivative at p in the direction v means that the curve $(p + tv, f(p + tv))$ in $\text{graph}(f)$ through $(p, f(p))$, which lies over the line in the (x, y) plane through the vector v , has a velocity vector at $t=0$. Thus even if we ask more of our function f , for instance if we ask that it have directional derivatives in every direction, we almost get condition (i) above but not necessarily condition (ii). For instance, if we define $f(x, y) = r\theta$, where $0 \leq \theta < \pi$ and $-\infty < r < \infty$, we get a function which has partials at $(0, 0)$, and even has directional derivatives at $(0, 0)$ in every direction, but which is not even continuous at $(0, 0)$, and such that the velocity vectors at $(0, 0)$ to curves in $\text{graph}(f)$ in different directions do not all lie in the same plane. This graph is made up of lines through the origin. Start with the x -axis as the first line and nail it down at the origin but don't nail it anywhere else. Then take hold of the line at the point $(1, 0)$ and walk with it counterclockwise around the unit circle in the upper half of the x, y plane as you lift up on the line. The line still passes through the origin, but has height θ over the point at angle θ on the upper half of the unit circle.

This function is not continuous but has directional derivatives in every direction; the corresponding velocity vectors to curves in the graph do not lie in a plane. You can make a variation on this example which is continuous by starting out the same but when you get to angle $\pi/4$, start letting the line down again so that at angle $\pi/2$ it becomes the y axis. Then do the same thing over again in the second quadrant. This gives a function whose partials are both zero at $(0,0)$, but which has non zero directional derivatives in every other direction. Neither of these examples satisfies condition (ii) and hence neither is differentiable.

The upshot of all this is that we must define carefully what it means for $f(x,y)$ to be differentiable at p , essentially by requiring that the graph have a tangent plane at $(p,f(p))$. It will then follow that every curve in the graph through $(p,f(p))$ has a tangent line there and that all these lines lie in the tangent plane. The basic facts are the following:

If $f: \mathbb{R}^k \rightarrow \mathbb{R}^n$ is "differentiable at p ", then it has partial derivatives at p , The matrix $[f'(p)]$ whose columns are the vector partials, is called the Jacobian matrix of f at p . f will also have directional derivatives in every direction v in \mathbb{R}^k , and in fact $D_v f(p)$ can be computed by multiplying the column vector v by the Jacobian matrix; i.e. we have the formula:

$D_v f(p) = [f'(p)][v]$, for every v in \mathbb{R}^k , where $[f'(p)]$ is the matrix whose columns are the vector partials of f .

The affine linear function $A(x) = f(p) + [f'(p)][x-p]$, is "tangent to f at p ", in the sense that its graph is the unique k -plane in \mathbb{R}^{k+n} which is tangent at $(p,f(p))$ to the graph of f . For values of x near p , this is the best affine linear function to use for approximating f .

It still remains for us to give a precise mathematical definition of the statement "f is differentiable at p " in a way that lives up to our intuition that it should mean that the graph of f has a tangent space at $(p,f(p))$. We do this as follows:

1) Define a function $\phi(t)$ to be "tangent to zero" (at $t=0$), if the ratio $\|\phi(t)\| / \|t\| \rightarrow 0$, as $\|t\| \rightarrow 0$. By looking at a picture in two variables we can see that this means that the graph is tangent to the (x,y) axis, which is the graph of the zero function.

2) Next define two functions f, g to be tangent to each other at p , if the difference $f(p+t) - g(p+t)$ is tangent to zero. This means essentially that $f(p) = g(p)$, and their graphs are tangent to each other at the common point $(p,f(p))$.

4

3) Last of all define a function f to be "differentiable at p " if f is tangent to some affine linear function at p , i.e. if and only if there is some homogeneous linear function $L(t)$, such that $f(x)$ is tangent to the affine linear function $A(x) = f(p) + L(x-p)$ at p . This says that for some linear function L , the ratio $\|f(x)-f(p)-L(x-p)\|/\|x-p\| \rightarrow 0$, as $x \rightarrow p$. If we use the symbol t with $t = x-p$, then $x = p+t$, and the statement says that $\|f(p+t)-f(p)-L(t)\|/\|t\| \rightarrow 0$, as $t \rightarrow 0$.

We could have given definition 3) without any of the previous definitions or any of the earlier discussion, but I hope this way of doing things has made it more understandable. Be aware however, that when it comes to the question of memorizing the definition of "differentiable", it is all contained in this sentence:

Definition: $f: \mathbb{R}^k \rightarrow \mathbb{R}^n$ is "differentiable at p " if and only if there is a (homogeneous) linear function $L: \mathbb{R}^k \rightarrow \mathbb{R}^n$ such that the ratio $\|f(p+t)-f(p)-L(t)\|/\|t\| \rightarrow 0$, as $t \rightarrow 0$.

If this is the case then L is called the (total) differential of f at p . L is denoted by the symbol $d_p f$. The function $A(x) = f(p) + L(x-p) = f(p) + (d_p f)(x-p)$, is called the best affine approximation to f at p . The graph of A is the tangent space to the graph of f at $(p, f(p))$. In analogy with the notation Δy from one variable, we can define $\Delta_p f(t) = f(p+t)-f(p)$.

Then f is differentiable at p if there is a linear map $d_p f(t)$ which is tangent to $\Delta_p f(t)$ at $t=0$.

We want to prove a few things:

Theorem:

1) If f is differentiable at p then f also has partial derivatives at p , and the matrix $[f'(p)]$ whose columns are the vector partials of f at p , is the matrix of the linear map $d_p f$, i.e. $[L] = [d_p f] = [f'(p)]$.

2) If f is differentiable at p then f also has directional derivatives $D_v f(p)$ at p , in every direction v in \mathbb{R}^k . Moreover $D_v f(p) = d_p f(v) = [f'(p)][v]$; i.e. the directional derivative in the direction v is obtained simply by evaluating the differential on v , equivalently by multiplying v by the matrix of partials.

3) If f is only known to have partial derivatives, and if it has them not only at p but at all points near p , and if furthermore those partials are functions which are continuous at p , then in fact f is actually differentiable at p also.

5

Remark: Part 3) of this theorem says that although simply having partials does not guarantee having a derivative, having continuous partials does guarantee it. Since all the functions you will ever meet are likely to have continuous partials, the distinction between having partials and having a derivative does not have a lot of practical consequences. The lengthy discussion we gave pointing it out was intended to help you to better understand the meaning of the concept of derivative.