MA 1C RECITATION 04/12/12

1. Directional and Partial Derivatives

To begin, we will specialize our study from a general multivariable function ($f : \mathbf{R}^n \to \mathbf{R}^m$) to the case of a real-valued multivariable function $f : \mathbf{R}^n \to \mathbf{R}$. Such a function is sometimes called a *scalar field*. This is the simplest interesting case of a multivariable function, but it allows us to illustrate many of the features that are also present in a more general setting.

Since there is "calculus" in the title of this class, the first natural question is to ask is "how do you differentiate such a function"? Intuitively, differentation allows us to find the instantaneous rate of change as you move in the domain. This is easy enough in $\mathbf{R}$, since we can only move back and forth on the real line, but as you saw with limits, in $\mathbf{R}^2$ and higher dimensions, there are infinitely many ways to move! This is why we cannot just take a derivative in multivariable calculus. We must also choose a *direction*. This leads us to the following definition.

Definition 1.1. The **directional derivative** of a function f at a point $\mathbf{a}$ along a *unit vector* $\mathbf{u}$ is defined as

$$f'(\mathbf{a}; \mathbf{u}) := \lim_{h \to 0} \frac{f(\mathbf{a} + h\mathbf{u}) - f(\mathbf{a})}{h}.$$

Remark 1.2. You may also see the directional derivative denoted as $f_u(a)$.

Warning 1.3. Just like in the one-dimensional case, the limit may not exist. For instance, suppose that a function is not continuous at the point.

Warning 1.4. It is important that the vector $\mathbf{u}$ is normalized, or else you will get the wrong number.

Looking at the definition of directional derivative, we see that this is just a natural generalization of the one-dimensional derivative you know and love. Often the $\mathbf{u}$ that you choose will be the directions of your basis, such as $e_1 = (1, 0)$ and $e_2 = (0, 1)$ in $\mathbf{R}^2$. We use these so much that we give them a special name.

Definition 1.5. The **partial derivative** of $f(x_1, \ldots, x_n)$ at $\mathbf{a}$ with respect to x_k is given by

$$\frac{\partial f}{\partial x_k}(\mathbf{a}) := f'(\mathbf{a}, e_k).$$

Remark 1.6. You can find more about the partial derivative on Apostol page 254.

Note that $\frac{\partial f}{\partial x_k}$ is again a multivariate function, so we can differentiate this with respect to another variable x_ℓ to get a function $\frac{\partial}{\partial x_\ell} \frac{\partial f}{\partial x_k} = \frac{\partial^2 f}{\partial x_\ell \partial x_k}$, which is called a *mixed partial derivative*.

Date: April 12, 2012.

Warning 1.7. Note that the order matters! It is *not* always true that the equality

$$\frac{\partial^2 f}{\partial x_\ell \partial x_k} = \frac{\partial^2 f}{\partial x_k \partial x_\ell}$$

holds. However, in many "nice" contexts, this equality does hold. For instance, it is a basic theorem that if all partial derivatives exist and *are continuous* (important!) on a domain D, then (mixed) partial derivatives commute on that domain.

1.1. **How to calculate partial derivatives.** In practice, computing partial derivatives is very easy. For instance, to compute $\partial f/\partial x_k$, just think of everything except for the variable x_k as constant and differentiate just like you would for one variable.

Warning 1.8. If you want to do this one the homework, however, you need to cite Theorem 8.3 in Apostol or give an explanation of why this technique is valid. (This is worth thinking about, if only to make sure you understand everything that's been presented so far.)

Let's start with a simple example.

Example 1. Let's compute the partial derivatives for the function $f(x, y) = x^2 - 2xy + y^2$. Let's first do it the long way—no shortcuts—along x, just from the definitions.

We have

$$\begin{aligned}\frac{\partial f}{\partial x} &= f'((x, y); (1, 0)) \\ &= \lim_{h\to 0} \frac{((x+h)^2 - 2(x+h)y + y^2) - (x^2 - 2x + y^2)}{h} \\ &= \lim_{h\to 0} \frac{(x+h)^2 - x^2}{h} + \lim_{h\to 0} \frac{2(x+h)y - 2xy}{h} \\ &= 2x - 2y,\end{aligned}$$

where the last step is from the usual power rule for a one-dimensional derivative. Note in particular $\frac{\partial f}{\partial x}$ is a function in two variables, even though we have not explicitly written it as such.

Note that f is symmetric with respect to x and y, so by our shortcut technique, we easily compute

$$\frac{\partial f}{\partial y} = 2y - 2x.$$

Remark 1.9. We should probably write $\frac{\partial f}{\partial x}(x, y)$, but this notation is cumbersome, especially once you work in more than two variables. Just make sure that you under what the shorthand $\frac{\partial f}{\partial x}$ stands for.

2. Total Derivatives

While taking directional derivatives is tremendously useful, one nice thing about the derivative in one variable is that it captured *all* of the information about the instantaneous rate of change in one simple object. However, there are infinitely many directions along a multivariable function. Is there any hope of doing getting such an object in multivariable calculus? Is there a way to put this seemingly infinite amount of information into one convenient package?

It turns out that the answer is yes! We just need to take a pinch of Ma 1a, a dash of Ma 1b, and stir it all in a big pot of Ma 1c to get the following object.

Definition 2.1. The **total derivative** of f at $\mathbf{a}$ is a *linear transformation* $T_{\mathbf{a}}$ such that for all $\mathbf{v}$ in some ball about $\mathbf{a}$, we have

$$f(\mathbf{a}+\mathbf{v}) = f(\mathbf{a}) + T_{\mathbf{a}} + ||\mathbf{v}||E(\mathbf{a},\mathbf{v})$$

where E is a function such that $E(\mathbf{a},\mathbf{v}) \to 0$ as $||\mathbf{v}|| \to 0$.

Warning 2.2. Once again, such a thing might not exist if your function f is not sufficiently "nice."

Let's step back and try and parse this definition. If you look carefully, you just see an shadow of the one-dimensional derivative in disguise. Essentially, our definition says that if we consider $(f(\mathbf{a}+\mathbf{v}) - f(\mathbf{a}))/||v||$ as $||\mathbf{v}|| \to 0$, there exists a limit, which is the linear transformation $T_{\mathbf{a}}$.

The way I like to think about the total derivative $T_{\mathbf{a}}$ is that it is machine that eats vectors and spits out a number that indicates how f is changing in the direction of that vector. It's like a guru that knows everything about how a function f changes. If it can be found by a directional derivative, it's something that Swami Total Derivative already knows.

It's nice to know that such a thing can exist. In general, we know that a function can have directional derivatives in certain directions, but not in other directions. One useful consequence is that if a total derivative exists at a point, all directional derivatives also exist at that point. Indeed, for a direction $\mathbf{u}$, we have

$$f'(\mathbf{a};\mathbf{u}) = T_{\mathbf{a}}(\mathbf{u}).$$

3. Common Pitfalls

We've developed a lot of machinery thus far. It can be a lot to absorb all at once if it's the first time you've seen it, but it will all become second nature once you get more experience with it. However, there are some common errors and misconceptions in learning this stuff that I would like to address.

- A function can have all its partial derivatives, but not be differentiable! Consider the function

$$f(x,y) = \begin{cases} 0, & xy = 0 \\ \frac{1}{xy}, & \text{otherwise.} \end{cases}$$

  It doesn't have most directional derivatives, but its partial derivatives do exist!
- A function can have all its directional derivatives, but not be continuous! See Apostol, page 257 for an example. Note that this is different from the one-dimensional case, in which differentiability at a point applies continuity at a point. (Recall the proof of this. What fails in higher dimensions?)
- A function can be differentiable, but not have continuous partial derivatives! You will see this on your homework.

The general rule is **never assume anything is true unless you have a theorem, result, or proof to back it up**. This is part of the reason why we are tough on you and want you to explicitly cite theorems and results on your homework. Many things that seem "obvious" or "clear" to you may not even be true at all!

4. The Gradient

Here comes the first of the three grand objects in multivariable calculus. (The others are the divergence and the curl, which you will see later. It's not too far-fetched of a statement to say that multivariable calculus is the study of div, grad, and curl.)

Definition 4.1. The **gradient** of a function $f : \mathbf{R}^n \to \mathbf{R}$ at a point $\mathbf{a}$ is defined as

$$\nabla f(a) = \left(\frac{\partial f}{\partial x_1}(\mathbf{a}), \dots, \frac{\partial f}{\partial x_n}(\mathbf{a}) \right).$$

In particular, note that if f is *differentiable*, then the gradient of f is precisely the total derivative of f, written with respect to the standard basis.

Remark 4.2. Think of the gradient as a function from the set of n-variable functions to the set of n-vectors. You feed it a function, it spits out a vector. You can feed this vector a point to get information about partial derivatives in every direction at that point.

The gradient gives us access to many useful techniques, the first of which is the following useful result.

4.1. **The best way to compute directional derivatives.** If f is *differentiable*, then $f'(\mathbf{a};\mathbf{u}) = T_{\mathbf{a}}(\mathbf{u})$. Furthermore, we noted that $T_{\mathbf{a}}$ expressed with respect to the standard basis $\{e_i\}$ is precisely $\nabla f(\mathbf{a})$. Therefore, we see that

$$f'(\mathbf{a};\mathbf{u}) = \nabla f(\mathbf{a}) \cdot \mathbf{u},$$

where the $\cdot$ is the *dot product*, which is just multiplication taken coordinate-wise for the vectors.

5. When is a function differentiable?

We have seen that we have lots of nice properties that only work when a function is differentiable. Therefore, it is very important to have a way to tell if a function is differentiable. One of the most useful results for doing this is Apostol Theorem 8.7.

Theorem 5.1. *If all of the partial derivatives of f exist in some ball around $\boldsymbol{a}$ and the partials are continuous at $\boldsymbol{a}$, then f is differentiable at $\boldsymbol{a}$.*

Example 2. The function $f(x, y) = -\cos(xy)$ is differentiable at $(0, 0)$.

By Theorem 8.3, we can find the partial derivatives by taking the derivative with respect to one variable and assuming that the other variables are constant. We compute

$$\frac{\partial f}{\partial x} = y \sin(xy), \quad \frac{\partial f}{\partial y} = x \sin(xy).$$

By various theorems from last week, we know that these partial derivatives exist and are continuous at $(0, 0)$. (Show this!) Therefore, by Theorem 8.7, f is differentiable at $(0, 0)$.

Example 3. This is a "lite" version of your homework problem. Show that $f(x, y) = x^2 + y$ is differentiable at $(0, 0)$.

I won't solve this in the slickest way, because I want to give you an idea of how to argue for one of your homework problems whose setting isn't as simple as this one.

We want to find a linear transformation $T = T_{(0,0)}$ such that

$$f(v) = 0 + T(v) + ||v||E((0,0), v) \tag{$*$}$$

for all v such that $||v|| < \epsilon$—where we can control $\epsilon > 0$—and such that $E(v) := E((0,0), v) \to 0$ as $v \to (0,0)$. This is difficult in general, so we want to use Theorem 8.7, which reduces our problem to just finding some T. If the partial derivatives exist, they are a good guess for what a T might look like. (This is because if they all exist and are continuous at $(0,0)$, then T is just the gradient.) Even better, if you can draw the function (e.g. if it's in a plane), then the $T(v)$ above is just the tangent plane, which gives you an even better idea of what the answer might be.

We can indeed draw our function in our case, and it seems like $T = (0,1)$ might be something that works. (How'd I find this? I sort of cheated. I used the gradient and evaluated at $(0,0)$.) Let's test it out. Substituting $T = (0,1)$ into the equation $(*)$, we obtain

$$v_1^2 + v_2 = (0,1) \cdot (v_1, v_2) + \sqrt{v_1^2 + v_2^2}E(v),$$

and with some algebra, we find that

$$E(v) = \frac{v_1^2}{\sqrt{v_1^2 + v_2^2}}.$$

Okay, so we have a plausible argument and a candidate solution. Let's prove that these do work.

We observe that this value of T does indeed satisfy the argument in a ball around $(0,0)$ and that E is well-defined in such a ball. (Check!) It remains to show that $E(v) \to 0$ as $v \to (0,0)$. Let $\epsilon > 0$ and set $\delta = \epsilon$. Therefore, if $||v|| < \delta$, then $|v_1| < \delta = \epsilon$. We have

$$|E(v)| = \left| \frac{v_1^2}{\sqrt{v_1^2 + v_2^2}} \right| \leq |v_1| < \epsilon,$$

proving our claim.

Example 4. Find a differentiable function whose maximal directional derivative at $(3,5)$ is equal to 1 in the direction $(1,0)$.

This kind of problem, which you also have on your homework, is a preview of an important concept which we will study later.

Recall that the directional derivative in the direction $(1,0)$ is just $\nabla f(3,5) \cdot (1,0)$. Therefore, we need to find a function f such that

$$\frac{\partial f}{\partial x}(3,5) = 1. \tag{1}$$

Now we need to find something that satisfies this "maximal" condition. Now, for $(1,0)$ to be the direction of maximal increase, $(1,0)$ must be the vector that maximizes $\nabla f(3,5) \cdot \mathbf{v}$. However, recall the following property of the dot product:

$$|\mathbf{a} \cdot \mathbf{b}| = ||\mathbf{a}||||\mathbf{b}|| \cos(\theta),$$

where θ is the angle between the vectors. Since $\cos(\theta)$ is maximized when $\cos(\theta) = 1$, we note that the dot product is maximized when $\theta = 0$, that is, when $\mathbf{a}$ and $\mathbf{b}$ lie on the same line. Thus, if we find a function such that

$$\nabla f(3,5) = (1,0) \tag{2}$$

we are done.

There are many functions that satisfy criteria (1) and (2). For instance, $f(x,y) = (x^3/3 - 5x) + (y-5)^2$ is one possible solution.