Ma1c Analytic Recitation 4/12/12

Alden Walker

1 Topology

Most of your homework questions last week were all about proving things about open sets. Really, you were studying $\mathbb{R}^n$ with its *standard topology*. A *topological space* is a pair $(X, \mathcal{T})$ where X is a set and $\mathcal{T}$ is a collection of subsets of X, which are called the open sets of X. The open sets must satisfy the following:

- X and $\varnothing$ are in $\mathcal{T}$.
- A finite interesection of sets in $\mathcal{T}$ is in $\mathcal{T}$ (a finite intersection of open sets is open).
- A arbitrary union of open sets is open

A example of a topological space is $(\mathbb{R}^n, \mathcal{T})$, where $\mathcal{T}$ is all unions of open balls. This is the standard topology that you are used to and you used in all of your homework problems.

There are other topologies on $\mathbb{R}^n$, however. Consider the space $(\mathbb{R}, \mathcal{T})$, where $\mathcal{T}$ is all complements of finite sets. This is called (I think) the finite complement topology. An example of a set in $\mathcal{T}$ is $\mathbb{R} \setminus \{0, 1, 2, 3\}$. This topology has the interesting property that every sequence of distinct values converges to every point. Why is this so? Given any sequence and point and open set about the point, there are only a finite number of points outside the open set. Therefore, there is some natural number past which the sequence must lie entirely in the open set.

The general definition of continuity is in the language of topology, also. It says that a function is continuous if the preimage of any open set is open. You can check that this is equivalent to the usual definition.

2 Homework Note

On problem 8.14.8, it gives you a hint to use problem 7(d). You may use the result of exercise 7(d) without proof (but you should say "here I am using the result of 7(d)").

3 Directional Derivatives

In all of the following, we will consider a function $f : \mathbb{R}^n \to \mathbb{R}$, i.e. a scalar field. We restrict the range to $\mathbb{R}$ because this case is simpler and there are more tricks you case use to get a handle on what's going on.

If the domain of the function is $\mathbb{R}$, then you know how to differentiate the function, but what does that mean in higher dimensions? To differentiate, you find, intuitively, the instantaneous change in the function when you move in the domain. In $\mathbb{R}$, you don't have to think about which direction you are moving, but in $\mathbb{R}^n$, there are lots of directions! The easiest way to generalize the notion of a one-dimensional derivative is to use a directional derivative:

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

Where $\mathbf{y}$ is a unit vector! When the limit exists. Notice this is really just the one-dimensional derivative of the one-dimensional function $g : \mathbb{R} \to \mathbb{R}$ defined by $g(t) = f(\mathbf{a} + t\mathbf{y})$. The most common directions that you want to know about are the basis directions. Thus, we call $f'(\mathbf{a}, \mathbf{e}_k)$ the partial derivative of f with respect to x_k. It is also denoted by $\frac{\partial f}{\partial x_k}$ (and a few other things you can find on page 254). You can, of course, take the partial with respect to x_1 and then the partial of *that* with respect to x_2, etc. This would be denoted $\frac{\partial^2 f}{\partial x_2 \partial x_1}$ and is called a mixed partial derivative. Note the order! It matters (in general—often it does not).

3.1 How to Compute Partials

Basically, to compute $\partial f/\partial x_k$, you just think of everything except x_k as constant and differentiate as you would if it were just a function of x_k. If you do this on your homework, however, you need to cite Theorem 8.3 in the book, or give a short explanation of why this is true! (Think about it).

3.2 Example

Let's compute the partials of $f(x, y) = x^4 + y^4 - 4x^2y^2$. We'll do the first straight from the definition.

$$\begin{aligned}\frac{\partial f}{\partial x} &= f'((x,y);(1,0)) \\ &= \lim_{h\to 0} \frac{((x+h)^4 + y^4 - 4(x+h)^2y^2) - x^4 + y^4 - 4x^2y^2}{h} \\ &= \lim_{h\to 0} \frac{(x+h)^4 - x^4}{h} + \lim_{h\to 0} \frac{4(x+h)^2y^2 - 4x^2y^2}{h}\end{aligned}$$

The last we recognize as simply a one-dimensional derivative, and we use the power rule to get $\frac{\partial f}{\partial x} = 4x^3 - 8xy^2$. Notice that even though it's not explicitly written, $\frac{\partial f}{\partial x}$ is a function of x and y, just like with the one-dimensional derivative.

Since the function looks identical in x and y, the other partial is the same, with x and y flipped.

4 Total Derivatives, etc

The last section covered how to find directional derivatives, but what do we mean by "the" derivative of a function on $\mathbb{R}^n$? The total derivative of f at $\mathbf{a}$ is a linear transformation $T_{\mathbf{a}}$ such that

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

For all $\mathbf{v}$ in some ball about $\mathbf{a}$, and where $E(\mathbf{a}, \mathbf{v}) \to 0$ as $\|\mathbf{v}\| \to 0$. Basically what this means is that if you take $(f(\mathbf{a}+\mathbf{v}) - f(\mathbf{a}))/\|\mathbf{v}\|$ as $\|\mathbf{v}\| \to 0$, then it has a limit, and that limit is given by $T_{\mathbf{a}}$. A good way to think about $T_{\mathbf{a}}$, which will generalize when we good at functions to $\mathbb{R}^m$, is that if you give $T_{\mathbf{a}}$ a vector, it spits out a number which tells you how f is changing in that direction. Essentially it encodes all the directional derivative information in one package.

4.1 Differentiable $\Rightarrow$ Directional Derivatives

It is certainly possible for a function to have some directional derivatives but not others. However, the existense of a total derivative guarantees the existence of all directional derivatives; in fact, $f'(\mathbf{a}; \mathbf{v}) = T_{\mathbf{a}}(\mathbf{v})$.

4.2 Warnings

- It is possible for a function to have all its partial derivatives but not be differentiable!. For example, consider the function which is zero on the axes and is defined to be $1/(xy)$ elsewhere. Then it clearly doesn't have most directional derivatives, but its partials do exist. I just made this example up so that it was very clear, but there are other, more subtle examples, like the function $\frac{xy^2}{x^2+y^4}$ if $x \neq 0$, and 0 if $x = 0$.

- It is possible for a function to have all its directional derivatives and not be continuous! See the example on page 257. However, if a funtion is differentiable (has a total derivative), then it is continuous.

- There are many warnings similar to this; on your homework you will prove that just because it is differentiable doesn't mean that the partials are continuous. Never assume that anything is true because it "seems like it should be."

4.3 Gradient

The gradient of a function f at a point $\mathbf{a}$, which is denoted $\nabla f(\mathbf{a})$ (often just ∇f to denote the function from $\mathbb{R}^n \to \mathbb{R}^n$), is defined to be

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

Notice that when f is differentiable, the gradient of f *is* the total derivative of f, written in the standard basis.

4.3.1 The Best Way to Compute Directional Derivatives

If f is differentiable, then we noted above that $f'(\mathbf{a}; \mathbf{y}) = T_{\mathbf{a}}(\mathbf{y})$, and we also noted that $T_{\mathbf{a}}$, expressed in the standard basis, is $\nabla f(\mathbf{a})$. Therefore, we see that $f'(\mathbf{a}; \mathbf{y}) = \nabla f(\mathbf{a}) \cdot \mathbf{y}$.

4.4 How to Tell if a Function is Differentiable

We have all these things that are useful if a function is differentiable, but no good way to tell if a function is differentiable. Here is a good condition (Theorem 8.7): If all the partials of f exist in some ball about $\mathbf{a}$, and the partials are continuous at $\mathbf{a}$, then f is differentiable at $\mathbf{a}$.

4.5 Example

Let's show that $f(x, y) = \sin(xy)$ is continuous at $(0, 0)$. First we note that by Theorem 8.3, we can find the partials by taking the other variable to be constant. Then we find that $\partial f/\partial x = y \sin(xy)$ exists and is continuous at $(0, 0)$ (by the various theorems from last week). Similarly, $\partial f/\partial y = x \sin(xy)$. Therefore, f is differentiable at $(0, 0)$. We have already computed the gradient at zero, but it's the same everywhere; it is $\nabla f = (y \sin(xy), x \sin(xy))$.

Let's compute the directional derivative at $(\pi/2, \pi/2)$ in the direction $(1, 1)$. Here is something tricky—we need to use a unit vector to compute the directional derivative! Thus, we compute $\nabla f(\pi/2, \pi/2) \cdot (1/\sqrt{2}, 1/\sqrt{2}) = (1/\sqrt{2})((\pi/2) \sin(\pi/4) + (\pi/2) \sin(\pi/4)) = \pi/2$. A valuable piece of information! Well, it could be if the function came from somewhere....

4.6 Example

You have to do a more complicated version of this on your homework, but let's prove that $f(x, y) = x^2 + y$ is differentiable at $(0, 0)$.

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

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

For all $\mathbf{v}$ such that $\|\mathbf{v}\| < \epsilon$, where we get to set ϵ, and such that $E(\mathbf{v}) = E((0, 0), \mathbf{v}) \to 0$ as $\mathbf{v} \to (0, 0)$.

So this is the hard part and the reason why you pretty much always want to use Theorem 8.7: you just have to produce T. If the partials exist, they are a good guess about what it should be. If you can picture the function (it's a function in the plane), then the tangent plane should be what you get if you ignore the last term (with E).

Anyway we guess that $T = (0, 1)$. Then we substitute in from the above equality to find E and see if it works:

$$\begin{aligned} f(\mathbf{v}) &= T(\mathbf{v}) + \|\mathbf{v}\| E(\mathbf{v}) \\ v_1^2 + v_2 &= (0, 1) \cdot (v_1, v_2) + \sqrt{v_1^2 + v_2^2} E(\mathbf{v}) \\ \frac{v_1^2}{\sqrt{v_1^2 + v_2^2}} &= E(\mathbf{v}) \end{aligned}$$

Where in the second line, technically $\mathbf{v}$ should be a column vector, but I replaced it just with the dot product.

Ok let's go through the proof that this works. First, this is certainly true in some ball about $(0,0)$, since that definition of E works. Second, we must verify that $E \to 0$ as $\mathbf{v} \to (0,0)$. We will write the proof of this claim out rigorously: given $\epsilon > 0$, set $\delta = \epsilon$. Then we note that if $\|\mathbf{v}\| < \delta$, then clearly $|v_1| < \delta = \epsilon$. We also note that $E(\mathbf{v}) = \frac{v_1^2}{\sqrt{v_1^2+v_2^2}} \leq |v_1|$, and therefore $E(\mathbf{v}) < \epsilon$, and we have proved our claim.

4.7 Example

Find a differentiable function whose maximal directional derivative at $(2,3)$ is equal to 1 in the direction $(1,0)$.

This is something that you will think about in the next sections, but it pertains to this homework in that it involves solving for things involving the gradient. Notice that since the directional derivative in the direction $(1,0)$ is $\nabla f(2,3) \cdot (1,0)$, to get this equal to 1 we just have to get $\frac{\partial}{\partial x}(2,3) = 1$. There are lots of ways we could do this.

Let's think about the other condition. To be the direction of maximum increase, $(1,0)$ would need to be the vector which maximizes $\nabla f(2,3) \cdot \mathbf{v}$. But notice that since $|\mathbf{a} \cdot \mathbf{b}| = \|\mathbf{a}\|\|\mathbf{b}\| \cos(\theta)$, where θ is the "angle" between the vectors, we maximize by setting $\mathbf{v}$ equal to the gradient (or we just notice that the dot product is maximized when the vectors are equal). That is, $\nabla f(2,3) = (1,0)$.

There are many functions which satisfy these criteria that we could give; one example is $f(x,y) = (x^2/2 - x) + (y-3)^2$.

4.8 Example

Let's do problem 8.14.10:

4.8.1 (a)

If f is differentiable inside an n-ball $B(\mathbf{a})$, and $\nabla f(\mathbf{x}) = 0$ for all $\mathbf{x}$ in $B(\mathbf{a})$, we'll show that f is constant on $B(\mathbf{a})$.

Ok suppose that we have two points $\mathbf{x}$ and $\mathbf{y}$ in $B(\mathbf{a})$ such that $f(\mathbf{x}) > f(\mathbf{y})$ (WLOG the inequality is that way). Define a new function $g : [0,1] \to \mathbb{R}$ by $g(t) = f((1-t)\mathbf{x} + t\mathbf{y})$. Note that $g(0) > g(1)$, so we can apply the Mean Value Theorem to say that there is some point $s \in [0,1]$ such that $g'(s) = \frac{g(1)-g(0)}{1} < 0$.

However, by Theorem 8.3, we see that $g'(s) = f'(\mathbf{x}+s(\mathbf{y}-\mathbf{x}); \mathbf{y}-\mathbf{x}) = \nabla f(\mathbf{x}+s(\mathbf{y}-\mathbf{x})) \cdot (\mathbf{y}-\mathbf{x})$. Therefore $\nabla f(\mathbf{x}+s(\mathbf{y}-\mathbf{x})) \cdot (\mathbf{y}-\mathbf{x}) < 0$, which is a contradiction because we must have $\nabla f(\mathbf{x}+s(\mathbf{y}-\mathbf{x})) = 0$. Therefore we must have been wrong in our supposition that we could produce two points with unequal function values, and thus f is constant.

4.8.2 (b)

If $f(\mathbf{x}) \leq f(\mathbf{a})$ for all $\mathbf{x} \in B(\mathbf{a})$, show that $\nabla f(\mathbf{a}) = 0$.

Take any unit vector $\mathbf{v}$. Then we see that

$$\nabla f(\mathbf{a}) \cdot v = \lim_{h \to 0} \frac{f(\mathbf{a}+h\mathbf{v}) - f(\mathbf{a})}{h} \leq 0$$

Where the inequality occurs because of our assumption. Now take $\mathbf{v} = \frac{1}{\|\nabla f(\mathbf{a})\|} \nabla f(\mathbf{a})$. Rewriting the above line,

$$0 \geq \frac{1}{\|\nabla f(\mathbf{a})\|} \nabla f(\mathbf{a}) \cdot \nabla f(\mathbf{a}) = \|\nabla f(\mathbf{a})\|$$

However, since $\|\cdot\|$ is always non-negative, we must have $\|\nabla f(\mathbf{a})\| = 0$, and thus $\nabla f(\mathbf{a}) = 0$, as desired.