Math 1c	TA: Padraic Bartlett
Recitation 2: The Derivative	
Week 2	*Caltech 2011*

1 Random Question

Is it possible to find a collection of open balls $\{B_{\mathbf{x}_i}(r_i)\}_{i=1}^{\infty}$ in $\mathbb{R}^n$ such that

- $\bigcup_{i=1}^{\infty} B_{\mathbf{x}_i}(r_i) \supseteq \mathbb{Q}^n$, and
- $\sum_{i=1}^{\infty} \text{volume}\left(B_{\mathbf{x}_i}(r_i)\right) < 1$?

2 The Derivative

2.1 The directional derivative: definitions, theorems, examples.

This lecture is centered around defining the idea of the **derivative** in $\mathbb{R}^n$. There are a number of possible ways to do this! One way is to generalize the idea of "slope" from $\mathbb{R}^1$.

In other words: in $\mathbb{R}^1$, the derivative of a function $f : \mathbb{R} \to \mathbb{R}$ at some point a is the "slope" of the graph $f(x) = y$ at the point $(a, f(a))$. Analogously, we could define the **directional derivative** of a function $f : \mathbb{R}^n \to \mathbb{R}$ at some point $\mathbf{a}$, *along some direction* $\mathbf{v}$, as the "slope" of f at the point $\mathbf{a}$, as measured in the direction $\mathbf{v}$. More formally:

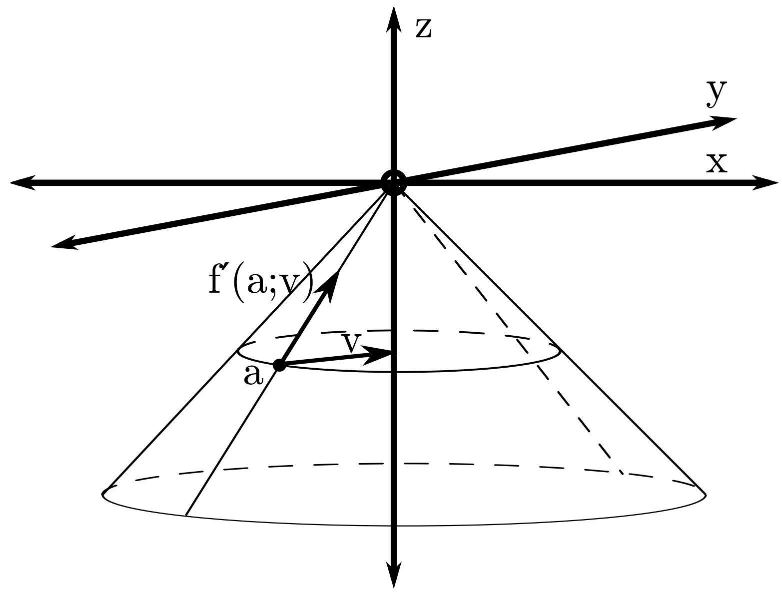
Definition. The **directional derivative** of a function $f : \mathbb{R}^n \to \mathbb{R}$ at some point $\mathbf{a}$ along some direction $\mathbf{v}$ is the limit

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

To illustrate what's going on here, consider the following example:

Question 1 *Consider the function* $f(x, y) = -\sqrt{x^2 + y^2}$. *What is the directional derivative of this function at the point* $(0, -1)$ *in the direction* $(0, 1)$*?*

Solution. First, to get a good idea of what's going on in this problem, we graph our function:



Visually, if we look at the point $(0,-1)$ and its slope in the direction $(0,1)$, we can see that it **should** be 1, just by examination. So, let's calculate, and see if our visual intuition matches our mathematical definition:

$$\begin{aligned}\lim_{h\to 0}\frac{f((0,-1)+h\cdot(0,1))-f((0,-1))}{h} &= \lim_{h\to 0}\frac{f((0,h-1))-f((0,-1))}{h\cdot\|(0,1)\|}\\ &= \lim_{h\to 0}\frac{(-\sqrt{0^2+(h-1)^2})-(-\sqrt{0^2+(-1)^2})}{h\cdot 1}\\ &= \lim_{h\to 0}\frac{-|h-1|+1}{h}\\ &= \lim_{h\to 0}\frac{h}{h}\\ &= 1,\end{aligned}$$

because for very small values of h, $-|h-1| = h-1$. So this matches our intuition!

Some of the most commonly-occuring directional derivatives are the **partial derivatives**, which we define below:

Definition. The **partial derivative** $\frac{\partial f}{\partial x_i}$ of a function $f:\mathbb{R}^n\to\mathbb{R}$ along its i-th coördinate at some point $\mathbf{a}$ is just the directional derivative $f'(\mathbf{a};\mathbf{e}_i)$: in other words, the limit

$$\lim_{h\to 0}\frac{f(\mathbf{a}+h\cdot\mathbf{e}_i)-f(\mathbf{a})}{h}.$$

Equivalently, it is just the derivative of f if we "hold all of f's other variables constant" – i.e. if we think of f as a single-variable function with variable x_i, and treat all of the other x_j's as constants. This method is markedly easier to work with, and is how we actually, say, *calculate* partial derivatives.

So: as it turns out, knowing these partial derivatives tells us exactly how to find *any* directional derivative! In particular, we have the following theorem:

Theorem 2 *The **directional derivative** of a function $f : \mathbb{R}^n \to \mathbb{R}$ at some point $\mathbf{a}$ along some direction $\mathbf{v}$ is given by the dot product of the gradient of f at $\mathbf{a}$,*

$$\nabla f\Big|_{\mathbf{a}} := \left(\frac{\partial f}{\partial x_1}(\mathbf{a}), \ldots, \frac{\partial f}{\partial x_n}(\mathbf{a})\right)$$

with $\mathbf{v}/||\mathbf{v}||$. In other words,

$$f'(\mathbf{a}; \mathbf{v}) := \nabla f\Big|_{\mathbf{a}} \cdot \frac{\mathbf{v}}{||\mathbf{v}||}$$

To illustrate the use of this theorem, return to our cone problem from earlier. There, we had $f(x,y) = -\sqrt{x^2+y^2}$; thus, if we hold y constant, we can see that

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

Similarly, by holding y constant, we have

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

Therefore, we know that the directional derivative of f at $(0,-1)$ in the direction $(0,1)$ is given by

$$\begin{aligned}\left(\frac{\partial f}{\partial x}(0,-1), \frac{\partial f}{\partial y}(0,-1)\right) \cdot (0,1) &= \left(\frac{-(0)}{\sqrt{0^2+(-1)^2}}, \frac{-(-1)}{\sqrt{0^2+(-1)^2}}\right) \cdot (0,1) \\ &= (0,1) \cdot (0,1) \\ &= 1,\end{aligned}$$

which matches our earlier answer.

2.2 The total derivative: definitions, theorems

So: as it turns out, the above notion is not the only way we have of thinking about derivatives! In addition to the geometric notion of "slope" in a given direction, we also had the more algebraic notion of a derivative being a "linear approximation" of a function in a given direction.

Specifically, for $\mathbb{R}^1$, the derivative $f'(a)$ was a constant such that the function $f(a) + xf'(a)$ was "very close" to $f(x)$ near a: i.e. it was a constant chosen such that the limit

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

(i.e. in the above limit, subtracting $f(a)+f'(a)\cdot h$ took away the "linear" part of f, leaving it with only (if f had a Taylor series $\sum c_i x^i$) terms that are quadratic or higher-order.)

Analogously, for $f:\mathbb{R}^n \to \mathbb{R}$, we can ask that the derivative be something similar! Specifically, consider the following definition:

Definition. The function $f:\mathbb{R}^n \to \mathbb{R}$ has a **total derivative** $T_{\mathbf{a}}$ at some point $\mathbf{a}$ if $f(\mathbf{a}) + T_{\mathbf{a}} \cdot (\mathbf{x})$ is a "linear approximation" of f at $\mathbf{a}$: i.e. if the limit

$$\lim_{||\mathbf{h}||\to 0} \frac{f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) - T_{\mathbf{a}} \cdot \mathbf{h}}{||\mathbf{h}||} = 0.$$

While this definition of the derivative has the advantage that it captures this idea of a "linear approximation" in a way that the directional derivative doesn't obviously do, it has the downside that it seems impossible to calculate! How can we find such a thing?

Well, as it turns out, with the directional derivative! In particular, we have the following theorem:

Theorem 3 *If $f:\mathbb{R}^n \to \mathbb{R}$ has a total derivative at the point $\supset$, then this total derivative is simply the gradient of f at $\mathbf{a}$: i.e.*

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

A quick consequence of the above theorem is that if f has a total derivative at some point $\mathbf{a}$, it has all of its directional derivatives at that point a. A question we could then ask is the following: does the converse hold? In other words, if a function f has all of its partial derivatives at some point, does it have a total derivative at that point?

As it turns out: no! Consider the following example:

Example. The function

$$f(x,y) = \begin{cases} \frac{y^3}{x^2+y^2}, & (x,y) \neq (0,0) \\ 0, & (x,y) = (0,0) \end{cases}$$

has partial derivatives $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$ defined on all of $\mathbb{R}^2$, and yet has no total derivative at $(0,0)$.

Solution. To find f's partial derivatives, we simply calculate and break things apart into cases. Specifically, for $\frac{\partial f}{\partial x}$, there are two possible situations we can find ourself in: either $y \neq 0$, or $y = 0$. In the first case, we have (by differentiating)

$$\frac{\partial f}{\partial x} = \frac{-2xy^3}{(x^2+y^2)^2}.$$

In the second, because $y = 0 \Rightarrow f(x,y) = 0$, it doesn't matter whether we're looking at $\frac{\partial f}{\partial x}(x,0)$ with $x \neq 0$, or $\frac{\partial f}{\partial x}(0,0)$; in either case

$$\frac{\partial f}{\partial x} = 0.$$

(Normally, we'd have to worry about $(0, 0$ as a special case, because our function is piece-wise defined; it's possible for the derivative to do something weird at the origin as a result of this. See the last example problem in this recitation for a situation where this happens!)

Similarly, for $\frac{\partial f}{\partial y}$, we have that whenever $x \neq 0$, we have (by differentiating)

$$\frac{\partial f}{\partial y} = \frac{3y^2}{x^2 + y^2} - \frac{2y^4}{(x^2 + y^2)^2}$$

and whenever $x = 0$, we have (as $f(0, y) = \frac{y^3}{y^2} = y$ for $y \neq 0$, and $f(0, 0) = 0 = y$ for $y = 0$)

$$\frac{\partial f}{\partial y} = 1.$$

So: we know that if our function *did* have a total derivative at $(0, 0)$, it would be given by the partials – i.e. that $T_{(0,0)}$, if it exists, must be $\left(\frac{\partial f}{\partial x}\Big|_{(0,0)}, \frac{\partial f}{\partial y}\Big|_{(0,0)} \right) = (0, 1)$.

However, when we examine the limit

$$\begin{aligned}
&\lim_{||(h_1,h_2)||\to 0} \frac{f((0,0) + (h_1, h_2)) - f(0,0) - T_{(0,0)} \cdot (h_1, h_2)}{||(h_1, h_2)||} \\
&= \lim_{||(h_1,h_2)||\to 0} \frac{\frac{h_2^3}{h_1^2 + h_2^2} - 0 - (0,1) \cdot (h_1, h_2)}{||(h_1, h_2)||} \\
&= \lim_{||(h_1,h_2)||\to 0} \frac{\frac{h_2^3}{||(h_1,h_2)||^2} - h_2}{||(h_1, h_2)||} \\
&= \lim_{||(h_1,h_2)||\to 0} \frac{h_2(h_2^2 - ||(h_1, h_2)||^2)}{||(h_1, h_2)||^3},
\end{aligned}$$

we can see that along the line $h_1 = h_2$, we have

$$\begin{aligned}
&\lim_{||(h_1,h_2)||\to 0} \frac{h_2(h_2^2 - ||(h_1, h_2)||^2)}{||(h_1, h_2)||^3} \\
&= \lim_{h_2 \to 0} \frac{h_2(h_2^2 - (\sqrt{h_2^2 + h_2^2})^2)}{(\sqrt{h_2^2 + h_2^2})^3} \\
&= \lim_{h_2 \to 0} \frac{h_2(h_2^2 - 2h_2^2)}{(\sqrt{2h_2^2})^3} \\
&= \lim_{h_2 \to 0} \frac{-h_2^3}{(\sqrt{2})^3 \cdot |h_2|^3} \\
&= \lim_{h_2 \to 0} -\frac{h_2/|h_2|}{2\sqrt{2}},
\end{aligned}$$

which clearly doesn't have a limit (and definitely doesn't converge to 0!) as $h_2 \to 0$. So our function is not totally differentiable at $(0, 0)$.

What went wrong above? Well, while our function *did* have all of its partial derivatives, they weren't exactly the nicest partial derivatives you could hope for. In other words, if you look at $\frac{\partial f}{\partial x}$: on any values along the line $x = y$ where $x \neq 0$ we have

$$\frac{\partial f}{\partial x} = \frac{-2x \cdot x^3}{(x^2 + x^2)^2} = \frac{-2x^4}{4x^4} = -\frac{1}{2}.$$

However when $x = y = 0$, we have $\frac{\partial f}{\partial x}(0, 0) = 0$! So, this partial is not a **continuous** function on all of $\mathbb{R}^2$. As a result, its behavior at $(0, 0)$ – which we use to define $T_{(0,0)}$ – doesn't accurately represent its behavior *near* (0,0), which is why $T_{(0,0)}$ failed to be a good linear approximation to f near $(0, 0)$.

As it turns out, this is the *only* way in which we can have partials and yet not have a total derivative! Formally, we have the following theorem:

Theorem 4 *For a function $f : \mathbb{R}^n \to \mathbb{R}$, if all of f's partials exist in a neighborhood of* **a** *and are continuous at the point* **a**, *then f has a total derivative at* **a**.

The last question we ask in our recitation, then, is whether continuity of the partials is *necessary* for a function to have a total derivative. In other words: suppose we have a function $f : \mathbb{R}^n \to \mathbb{R}$ has all of its partials defined at some point **a**, but some of them are discontinuous at **a**. Can f still somehow have a total derivative at **a**?

The answer, as we discuss with our last example, is yes!

Example. The function

$$f(x, y) = \begin{cases} x^2y^2 \cdot \sin\left(\frac{1}{xy}\right), & x \neq 0 \text{ and } y \neq 0, \\ 0, & \text{otherwise,} \end{cases}$$

has $\frac{\partial f}{\partial x}$ discontinuous at every point of the form $(0, y)$ where $y \neq 0$, $\frac{\partial f}{\partial y}$ discontinuous at every point of the form $(x, 0)$ where $x \neq 0$, and yet has a total derivative everywhere.

Solution. As an aside, note that this function is a generalization of the map $x \mapsto x^2 \sin(1/x)$, which (from back in Ma1a / Ma8!) was an example of a function on $\mathbb{R}^1$ that was differentiable but not C^1.

We calculate f's partials, starting with $\frac{\partial f}{\partial x}$. If $y = 0$, we have $f(x, 0) = 0$; therefore, $\frac{\partial f}{\partial x} = 0$.

Otherwise, if $y \neq 0$, we have two situations: either we're finding $\frac{\partial f}{\partial x}$ at some point where $x \neq 0$, in which case we can just differentiate normally:

$$\frac{\partial f}{\partial x} = 2xy^2 \cdot \sin\left(\frac{1}{xy}\right) - y \cos\left(\frac{1}{xy}\right).$$

However, if we're in the situation where we want to find $\frac{\partial f}{\partial x}$ at a point where $y \neq 0, x = 0$, because our function is piece-wise defined, we have to use the definition of the derivative to find $\frac{\partial f}{\partial x}$:

$$\begin{aligned}\frac{\partial f}{\partial x} &= \lim_{h\to 0} \frac{f(y, 0+h) - f(y,0)}{h} \\ &= \lim_{h\to 0} \frac{h^2y^2 \cdot \sin\left(\frac{1}{xy}\right) - 0}{h} \\ &= \lim_{h\to 0} hy^2 \cdot \sin\left(\frac{1}{xy}\right) \\ &= 0.\end{aligned}$$

Similarly, for $\frac{\partial f}{\partial y}$, we can show that whenever both $x, y \neq 0$, we have

$$\frac{\partial f}{\partial y} = 2x^2y \cdot \sin\left(\frac{1}{xy}\right) - x\cos\left(\frac{1}{xy}\right),$$

and whenever either of x or y are zero, $\frac{\partial f}{\partial y} = 0$.

So: notice that $\frac{\partial f}{\partial x}$ is discontinuous at all of the points $(0, y)$ where $y \neq 0$: this is because its derivative at $(0, y)$ is 0, while the limit as x approaches zero of the function $2xy^2 \cdot \sin\left(\frac{1}{xy}\right) - y\cos\left(\frac{1}{xy}\right)$ does not exist (and in fact fluctuates rapidly between y and $-y$ as x approaches 0.) Similarly, we can see that $\frac{\partial f}{\partial y}$ is discontinuous at all of the points $(x, 0)$ where $x \neq 0$.

However, as it turns out, our function has a total derivative at all of these points! In specific, we have (for points of the form $(0, y)$) the limit

$$\begin{aligned}&\lim_{||(h_1,h_2)||\to 0} \frac{f((0,y)+(h_1,h_2)) - f(0,y) - T_{(0,y)} \cdot (h_1,h_2)}{||(h_1,h_2)||} \\ = &\lim_{||(h_1,h_2)||\to 0} \frac{(h_1^2 \cdot (y+h_2)^2)\sin\left(\frac{1}{h_1(y+h_2)}\right) - 0 - (0,0)\cdot(h_1,h_2)}{||(h_1,h_2)||}. \\ = &\lim_{||(h_1,h_2)||\to 0} \frac{(h_1^2 \cdot (y+h_2)^2)\sin\left(\frac{1}{h_1(y+h_2)}\right)}{||(h_1,h_2)||}.\end{aligned}$$

The fraction in this limit is bounded below by

$$\frac{-h_1^2}{||(h_1,h_2)||} \cdot (y+h_2^2) \geq -\frac{h_1^2}{|h_1|} \cdot (2y)^2 = -|h_1| \cdot (2y)^2,$$

and above by

$$\frac{h_1^2}{||(h_1,h_2)||} \cdot (y+h_2)^2 \leq \frac{h_1^2}{|h_1|} \cdot (2y)^2 = |h_1| \cdot (2y)^2,$$

for all values of $h_2 < y$. So: notice that $\lim_{(h_1,h_2)\to 0} |h_1| \cdot (2y)^2$ is clearly 0; therefore, as $(h_1, h_2) \to 0$, we know that both of these bounding limits converge to 0. Therefore, by the squeeze theorem, we know that our original limit must converge to 0 as well – and thus that our function has a total derivative at all points of the form $(0, y)$, where $y \neq 0$!

Similar logic shows that it has a total derivative at all points of the form $(x, 0)$ where $x \neq 0$, as well as at $(0, 0)$: then, because our partials *are* continuous on the open set $\{(x, y) : x \neq 0 \text{ and } y \neq 0\}$, we know that our function has a total derivative everywhere in $\mathbb{R}^2$! So we've proven our claim.