MA 200 - Lecture 2

1 Recap

1. Motivated multivariable calculus and logistics of the course.

2. Reviewed linear algebra. In particular, vector spaces, dimension, linear maps, matrices, determinants, inner products, norms, and matrix norms.

2 Review of topology of $\mathbb{R}^n$

The inner product norm induces a metric in the sense of metric spaces, i.e., $d(x,y) = d(y,x), d(x,z) \leq d(x,y) + d(y,z), d(x,y) \geq 0$ with equality iff $x = y$. Recall that once we have a way to talk about distances, we can define open balls $B_a(r)$. Once we have these, we can talk of open sets, interiors, and closed sets (including limits points and closure). We can also talk of convergence of sequences: $d(x_n, x) < \epsilon$ whenever $n > N$. Moreover, every closed and bounded set is compact (and vice-versa), i.e., every open cover has a finite subcover, and equivalently, every sequence has a convergent subsequence. Hopefully you did connected sets too: A set is connected iff it cannot be written as a disjoint union of two relatively open subsets. Moreover, a set in $\mathbb{R}$ is connected iff it is an interval.

Consider a function $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$. Then $\lim_{x \to x_0, x \in U} f(x) = L$ if $\|f(x) - L\| < \epsilon$ whenever $0 < \|x - x_0\| < \delta$ and $x \in U$. One can prove that this limit can be equivalently defined using sequences. A function is said to be continuous at x_0 if $\lim_{x \to x_0, x \in U} f(x) = f(x_0)$. If we restrict the domain of a continuous function, it still remains continuous (and likewise for limits). (But beware! if you enlarge the domain, the function might stop being continuous!) Equivalently, a function is continuous iff the inverse image of an open set is open.

Continuous functions take compact sets to compact sets. As a consequence, we have the extreme-value theorem. Also, uniform continuity. Continuous functions also take connected sets to connected sets. Continuous functions satisfy various properties (sum, product, quotient (these three hold for limits too), and composition). One can come up several examples (using continuity/limit laws. Polynomials for instance are continuous) and non-examples (using sequences along different paths). Another example of a continuous function:

$$f(x,y) = \begin{cases} \frac{x^2 y}{x^2+y^2} & \text{when } (x,y) \neq (0,0) \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

3 Derivatives

Recall that in one-variable calculus, $f'(a) = \lim_{h\to 0} \frac{f(a+h)-f(a)}{h}$. In more than one variable, unfortunately, this naive definition cannot work (because we cannot divide by a vector). A reasonable substitute is the notion of a directional derivative of a function $f : U \subset \mathbb{R}^n \to \mathbb{R}$ at an *interior* (why?) point $a \in U$ along a vector $\vec{v}$: $\nabla_v f(a) = \frac{df(a+tv)}{dt}|_{t=0}$. (Caution: When $v = 0$, the name "directional derivative" is somewhat of a misnomer. Moreover, since $\nabla_{cv} f(a) = c\nabla_v f(a)$, again this name is not completely appropriate.) Examples:

1. When $v = e_i$, the resulting directional derivative is called the partial derivative of f w.r.t x_i and is denoted as $\frac{\partial f}{\partial x_i}$. This quantity can be calculated easily using the various rules for one-variable differentiation. (Tidbit: The laws of nature are partial differential equations, i.e., equations involving partial derivatives.)

2. One can have directional derivatives at all points in all directions: Polynomials for instance (note that this is a one-variable question!)

3. It is certainly possible to have directional derivatives along some directions and not along some others:

$$f(x,y) = \begin{cases} \frac{xy}{x^2+y^2} & \text{when } (x,y) \neq (0,0) \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

   has directional derivatives at $(0,0)$ along e_1 for instance but not along $e_1 + e_2$.

4. It is possible to have directional derivatives along all directions at all points in a domain and yet fail to be even continuous!

$$f(x,y) = \begin{cases} \frac{x^2y}{x^4+y^2} & \text{when } (x,y) \neq (0,0) \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

The last example illustrates that the notion of a directional derivative is not a good enough notion. Indeed, differentiability is a "nicer" condition than continuity. It must imply continuity at the very least! (Another problem (albeit less important) with directional derivatives is that, apparently, we need to keep track of *infinitely* many numbers (one for each direction) at even a *single* point of the domain to understand how quickly the function changes at that point.) Recall that in one-variable calculus, $f'(a) = \lim_{h\to 0} \frac{f(a+h)-f(a)}{h}$. In more than one variable, unfortunately, this naive definition cannot work (because we cannot divide by a vector). A reasonable substitute is the notion of a directional derivative of a function $f : U \subset \mathbb{R}^n \to \mathbb{R}$ at an *interior* (why?) point $a \in U$ along a vector $\vec{v}$: $\nabla_v f(a) = \frac{df(a+tv)}{dt}|_{t=0}$. (Caution: When $v = 0$, the name "directional derivative" is somewhat of a misnomer. Moreover, since $\nabla_{cv} f(a) = c\nabla_v f(a)$, again this name is not completely appropriate.) Examples:

1. When $v = e_i$, the resulting directional derivative is called the partial derivative of f w.r.t x_i and is denoted as $\frac{\partial f}{\partial x_i}$. This quantity can be calculated easily using the various rules for one-variable differentiation. (Tidbit: The laws of nature are partial differential equations, i.e., equations involving partial derivatives.)

2. One can have directional derivatives at all points in all directions: Polynomials for instance (note that this is a one-variable question!)

3. It is certainly possible to have directional derivatives along some directions and not along some others:

$$f(x,y) = \begin{cases} \frac{xy}{x^2+y^2} & \text{when } (x,y) \neq (0,0) \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

   has directional derivatives at $(0,0)$ along e_1 for instance but not along $e_1 + e_2$.

4. It is possible to have directional derivatives along all directions at all points in a domain and yet fail to be even continuous!

$$f(x,y) = \begin{cases} \frac{x^2y}{x^4+y^2} & \text{when } (x,y) \neq (0,0) \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

The last example illustrates that the notion of a directional derivative is not a good enough notion. Indeed, differentiability is a "nicer" condition than continuity. It must imply continuity at the very least! (Another problem (albeit less important) with directional derivatives is that, apparently, we need to keep track of *infinitely* many numbers (one for each direction) at even a *single* point of the domain to understand how quickly the function changes at that point.)

Let us recall why differentiability implies continuity in one-variable calculus in the first place: $\left|\frac{f(a+h)-f(a)}{h} - f'(a)\right| < \epsilon$ when $0 < |h| < \delta$. Hence, $|f(a+h) - f(a) - f'(a)h| < \epsilon|h|$. Using the triangle inequality, $|f(a+h) - f(a)| < |h|(|f'(a)| + \epsilon)$. Using the squeeze rule, we are done.

In other words, the key point is the ability to *approximate* f well, i.e., $\lim_{h\to 0} \frac{f(a+h)-f(a)-f'(a)h}{h} = 0$. After all, one of the points of one-variable differential calculus is to approximate the curve by its tangent line. Likewise, we should expect to be able to approximate a function by its tangent plane, i.e., $f(a+h) \approx f(a) + L(a,h)$ where $L(a,h)$ is a point on a plane. (The same kind of a thing ought to hold even if f is vector-valued.) So surely, it is *linear* in h. (A plane passing through the origin is a *subspace* of $\mathbb{R}^n$. Hence, L is a linear map in h.) But differentiability is much more than mere continuity. So before we proceed to the definition, let's do a sanity check with the help of another example (which we also looked at in UM 102): $f(x,y) = ||x-y| - ||x| - |y|||$. This function is continuous at $(0,0)$ (composition of continuous ones). In no sense does a tangent plane exist at the origin. (The graph looks like a crumpled up piece of paper.)

Definition: Let $a \in U \subset \mathbb{R}^n$ be an interior point and $f : U \to \mathbb{R}^m$ be a function. It is differentiable at a if there exists a linear map $L : \mathbb{R}^n \to \mathbb{R}^m$ (sometimes called the total derivative or simply the derivative of f at a) such that $\lim_{h \to 0} \frac{\|f(a+h) - f(a) - L(h)\|}{\|h\|} = 0$. f is said to be differentiable on an open set U if it is differentiable at all points of U.

Remark: When S is an *arbitrary* set, many people define f to be differentiable on S if there is *some* open set U containing S such that f can be defined on U and is differentiable on U. This definition can be rather tricky to use. We will not bother with it (at least not right now).

Zeroethly, if f is differentiable at a, then L is unique: Indeed, if L_1, L_2 are two such maps, then $\|L_1(h) - L_2(h)\| = \|h\| \frac{\|L_1(h) - L_2(h)\|}{\|h\|} \leq \|h\| \frac{\|f(a+h) - f(a) - L_1(h)\|}{\|h\|} + \|h\| \frac{\|f(a+h) - f(a) - L_2(h)\|}{\|h\|}$. Thus, $\|(L_1 - L_2)(h)\| \leq \epsilon \|h\|$ as long as $\|h\| < \delta$. But $L_1 - L_2$ is linear and hence by scaling, this is true for all h! Since this inequality is true for all ϵ, $(L_1 - L_2)(h) = 0 \; \forall \; h$.

We are now faced with many questions: Does this notion of differentiability imply continuity? Can we now talk about a tangent plane? Can we hope to calculate $L(h)$ and is it related to the directional derivative? How can we *check* (come up with examples and non-examples) differentiability or the lack thereof in many cases? The answer to all of these questions is 'yes'.

We begin with the following proposition.

Proposition 3.1. *If $f : U \to \mathbb{R}$ is differentiable at a, all of its directional derivatives exist at a and $L(h) = \nabla_h f(a) = \frac{\partial f}{\partial x_1}(a) h_1 + \frac{\partial f}{\partial x_2}(a) h_2 + \ldots + \frac{\partial f}{\partial x_n}(a) h_n$.*

Defining (the derivative/gradient) ∇f as $\nabla f = (\frac{\partial f}{\partial x_1}, \ldots)$, we see that $L(h) = \langle \nabla f(a), h \rangle$. By the Cauchy-Schwarz inequality, $-\|\nabla f(a)\| \|v\| \leq \nabla_v f(a) \leq \|\nabla f(a)\| \|v\|$ with equality holding only when v is along/opposite to $\nabla f(a)$. Thus, $\nabla f(a)$ is the direction of steepest increase of f. (Hence the term, gradient.)