3 Differential Calculus

3.1 Limits & Continuity

A real sequence, or simply **sequence**, is a collection of numbers $a_1, a_2, a_3, \ldots$ that can be indexed $1, 2, 3, \ldots$. The sequence in the previous sentence can be abbreviated $\{a_n\}_{n=1}^{\infty}$, and is said to **converge** if there is a number a such that

$$\forall \epsilon > 0,\ \exists N \in \mathbb{N} \text{ such that } \forall n \geq N,\ |a_n - a| < \epsilon.$$

The number a, if it exists, is unique and is called the **limit** of the sequence $\{a_n\}_{n=1}^{\infty}$. To see why it is unique, suppose both a and a' were limits of the convergent sequence $\{a_n\}_{n=1}^{\infty}$. Then that would mean that for all $\epsilon > 0$, there are numbers N and N' such that $|a_n - a| < \epsilon$ for all $n \geq N$ and $|a_n - a'| < \epsilon$ for all $n \geq N'$. Then for $n \geq \max\{N, N'\}$,

$$|a - a'| = |(a - a_n) + (a_n - a')| \leq |a_n - a| + |a_n - a'| < \epsilon + \epsilon = 2\epsilon.$$

The first inequality in the centered statement is called the **triangle inequality**:

$$|a + b| \leq |a| + |b|,$$

which is true for all numbers a and b. The fact that $|a - b| = |b - a|$ is also put in use here. Since you can pick an ϵ arbitrarily small, this concludes the argument that $a = a'$. Therefore, the limit of a convergent sequence is unique. We often abbreviate the statement "the sequence $\{a_n\}_{n=1}^{\infty}$ converges to the limit a" as

$$\lim_{n \to \infty} a_n = a. \tag{9}$$

Exercise 20. Let $\{a_n\}_{n=1}^{\infty}$ and $\{b_n\}_{n=1}^{\infty}$ be convergent sequences with limits a and b respectively, and let c be a number. Convince yourself that the following statements are true: (a) $\lim_{n\to\infty} ca_n = ca$, (b) $\lim_{n\to\infty} a_n + b_n = a + b$, (c) $\lim_{n\to\infty} a_n - b_n = a - b$, (d) $\lim_{n\to\infty} a_n b_n = ab$, and (e) if $\forall n$, $b_n \neq 0$ and $b \neq 0$, then $\lim_{n\to\infty} a_n/b_n = a/b$.

If $\{a_n\}_{n=1}^{\infty}$ is a sequence then let $s_n = \sum_{k=1}^{n} a_k$. This gives rise to the sequence $\{s_n\}_{n=1}^{\infty}$ of **partial sums**. If this sequence converges to a limit s, then we say that the **series** $\sum_{n=1}^{\infty} a_n$ converges to the sum s. If the sequence of partial sums does not converge, then we say that the series diverges.

Exercise 21. If $|r| < 1$, a is a number and n a natural number, then prove that

$$a + ar + ar^2 + \cdots + ar^n = \frac{a(1 - r^n)}{1 - r}.$$

Then use this to show that the series $\sum_{n=0}^{\infty} ar^n$ converges to $a/(1 - r)$.

Exercise 22. This exercise describes the **multiplier effect** of spending. There are m shop-owners in Mali. A tourist enters Mali and spends \$10 at Ms 1's shop. Ms 1 takes 80% of her profit and spends it at Ms 2's shop; Ms 2 spends 80% of her profit at Mr 3's; ... and so on; Ms m spends 80% of her profit at Ms 1's, and this continues in a loop. For every dollar transaction at a Malian shop, 70 cents is the cost of the goods sold. What Malians do not spend at each others shops, they save at the Timbuktu Bank. What fraction of the \$10 spent by the tourist gets saved at the bank? What is the value of total purchases by Malians resulting from the tourist spending \$10 at Ms 1's shop?

Now, we want to capture the idea that a function $f : S \to \mathbb{R}$ (where S is an interval, could be $(-\infty, \infty)$) is "**continuous** at $x \in S$" if for all sequences $\{x_n\}_{i=1}^{\infty}$ that converge to x, the sequence $\{f(x_i)\}_{i=1}^{\infty}$ converges to $f(x)$. By the definition of convergence, this means that if $\{x_i\}_{i=1}^{\infty}$ converges to x, then

$$\forall \epsilon > 0, \exists N \text{ s.t. } \forall n \geq N, |f(x_n) - f(x)| < \epsilon.$$

But by the definition of $\{x_i\}_{i=1}^{\infty}$ converging to x, this is the same as saying

$$\forall \epsilon > 0, \exists \delta > 0 \text{ such that } y \in S \text{ and } |x - y| < \delta \text{ implies } |f(x) - f(y)| < \epsilon.$$

We say that f is "a continuous function" if it is continuous at every point in S.

Exercise 23. Note that the sum and product of two continuous functions are also continuous. Prove that the composition of two continuous functions is continuous.

There is a result in calculus (that we will not cover) that says that the series

$$\sum_{n=1}^{\infty} \frac{x^{n-1}}{(n-1)!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + ... =: f(x)$$

converges and that the derivative of the function f whose domain is the real numbers, can be found by differentiating the elements of the sum, term by term. Thus,

$$f'(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots \tag{10}$$

That's not strange. In fact, $f(x) = f'(x) = f''(x) = \cdots$ for this function, and we have a special name for it. We call such a function $f(x) = \exp(x)$, and it turns out that this function equals a number e (which happens to be irrational) raised to the power x. Thus $f(x) = e^x$. This function is strictly increasing and bijective if its range is defined to be only the positive numbers. (In fact, you can draw a graph of it to verify this.) Recall that bijective functions are invertible. The associated inverse function is called $\log x$.

3.2 The Derivative

The function $f : S \to \mathbb{R}$ is "**differentiable** at $x \in S$" if S is an open interval and $\exists a \in \mathbb{R}$ such that

$$\forall \epsilon > 0, \exists \delta > 0 \text{ such that } y \in S \text{ and } |x - y| < \delta \text{ implies } \left| \frac{f(x) - f(y)}{x - y} - a \right| < \epsilon.$$

It is a "differentiable function" if it is differentiable at every point in S.

Typically, the number a will depend on x, so we may as well write $a(x)$. If $a(x)$ is unique (which it is, and you can verify this), then $\{(x, a(x)) : x \in S\}$ is a function over $S \times \mathbb{R}$ whenever f is differentiable. In that case, we define the function $f' : S \to \mathbb{R}$, with $f'(x) = a(x)$, which we call the (first) **derivative** of f. The derivative of f', if it exists, is denoted f'', and is called the second derivative of f, and so on.

It is also important to know that we can define differentiability another way. If for all sequences $\{x_n\}_{n=1}^{\infty}$ such that $\lim_{n\to\infty} x_n = y$ and $x_n \neq y$ for all n we have

$$\lim_{n\to\infty} \frac{f(x_n) - f(y)}{x_n - y} = f'(y) \tag{11}$$

then we say that f is differentiable at y, where its derivative is $f'(y)$. Often, we abuse notation to write this statement as

$$\lim_{x\to y} \frac{f(x) - f(y)}{x - y} = f'(y). \tag{12}$$

In fact, I'll call this limit the "abusive limit," to be read as "limit as x reaches y..." Similarly, I'll use the notation $\lim_{x\to y^+}$ to mean that the limit holds for all sequences $\{x_i\}_{i=1}^{\infty} \subset (y, \infty)$ that converge to y, and $\lim_{x\to y^-}$ means that the limit holds for all sequences $\{x_i\}_{i=1}^{\infty} \subset (-\infty, y)$ that converge to y. Also as an abuse of notation, $\lim_{x\to+\infty}$ means that the sequence in consideration increases without bound while $\lim_{x\to-\infty}$ means that it decreases without bound.

Exercise 24. Convince yourself that the two definitions of differentiability are equivalent. That is, derive the second from the first, and the first from the second. (*Hint*: Write down the ϵ, δ definition of the limit in (11).) Also convince yourself that if a function is differentiable, then it is continuous. (*Hint*: Multiply the last expression in the ϵ, δ definition of differentiability by $|x - y|$.)

3.3 Properties of the Derivative

If $f : S \to \mathbb{R}$ and $g : S \to \mathbb{R}$ are differentiable at $y \in S$ and c is a number, then cf, $f + g$, $f - g$ and fg are all differentiable at y. Here, cf is the function defined by multiplying

$f(x)$ by c at all $x \in S$, $f + g$ is the function defined by adding $f(x)$ to $g(x)$ at all $x \in S$. Instead of adding, we subtract to define $f - g$ and multiply to define fg. If $g(x) \neq 0$ for all $x \in S$, then f/g, which is the function defined by dividing $f(x)$ by $g(x)$, is also differentiable. In fact, it is easy to show that

$$\begin{aligned} [cf]'(y) &= cf'(y), \\ [f+g]'(y) &= f'(y) + g'(y), \\ [f-g]'(y) &= f'(y) - g'(y) \end{aligned}$$

Now notice that

$$\frac{f(x)g(x) - f(y)g(y)}{x - y} = f(x)\frac{g(x) - g(y)}{x - y} + \frac{f(x) - f(y)}{x - y}g(y), \tag{13}$$

which is the main step in the proof of the **product rule**:

$$[fg]'(y) = f(y)g'(y) + f'(y)g(y). \tag{14}$$

In fact, all that one has to do is take the abusive limit on both sides of (13) and then use the fact that differentiable functions are continuous. Similarly, notice that

$$\frac{1/g(x) - 1/g(y)}{x - y} = -\frac{1}{g(x)g(y)}\frac{g(x) - g(y)}{x - y} \tag{15}$$

helps prove that $\left[\frac{1}{g}\right]'(y) = -\frac{g'(y)}{(g(y))^2}$. Take the abusive limit on both sides of (15) and combine this result with the product rule to get the **quotient rule**:

$$\left[\frac{f}{g}\right]'(y) = \frac{g(y)f'(y) - g'(y)f(y)}{(g(y))^2}. \tag{16}$$

Think about why we are allowed to take the abusive limit on both sides.

Finally, let $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ be two functions and assume the composition $f \circ g$ is defined on an open interval, S. Suppose that g is differentiable at $x \in S$ and that f is differentiable at $g(x)$. Then $f \circ g$ is differentiable at x, with derivative

$$[f \circ g]'(x) = f'(g(x))g'(x). \tag{17}$$

This fact is known as the **chain rule**, and the following argument shows why it is true. Since f is differentiable at $g(x)$, then there is an error term $r(y)$, implicitly defined for any $y \in S$ by

$$f(g(y)) - f(g(x)) = [f'(g(x)) + r(y)][g(y) - g(x)]; \tag{18}$$

this error term has limit 0 as $g(y) \to g(x)$. But by the definition of continuity, it has limit 0 as $y \to x$ as well. Now divide both sides of (18) by $y - x$ and take the abusive limit on both sides. On the left hand side you will get $[f \circ g]'(x)$. On the right hand side, the $r(y)$ term will vanish.

Exercise 25. Let $f : \mathbb{R} \to \mathbb{R}$ be a function defined by $f(x) = ax^n$ where $a \in \mathbb{R}$ and $n \in \mathbb{R}$. Find its 1st, 2nd and 3rd derivatives using the limits definition of the derivative.

3.4 The Derivative in Multiple Dimensions

Let $f : S \to \mathbb{R}$ be a function and $S \subset \mathbb{R}^n$. Suppose that for any $\epsilon > 0$ there exists $\delta > 0$ such that if $y \in S$, $||x - y|| < \delta$ implies that $|f(x) - f(y)| < \epsilon$ then f is said to be **continuous** at x. If the statement is true for every $x \in S$ then f is said to be a continuous function.

Similarly, let S_1, S_2, ..., S_n be open intervals; we can allow some or all of them to be $(-\infty, \infty)$. Then $S := S_1 \times ... \times S_n$ is a subset of $\mathbb{R}^n$, and we call S an **open box**. A function $f : S \to \mathbb{R}$ is said to be **differentiable** at $x \in S$ if for all $\epsilon > 0$ there is a $\delta > 0$ such that $y \in S$ and $||x - y|| < \delta$ implies

$$|f(x) - f(y) - a(x) \cdot (x - y)| < \epsilon ||x - y||,$$

for some vector $a(x)$ of size n. Akin to the one-dimensional case, the vector $a(x)$ is called the **derivative** of f at $x \in S$ and is unique for each x whenever it exists. If f is differentiable at all points in S then it is a differentiable function, and we can define the derivative of f to be the function $\nabla f : S \to \mathbb{R}^n$ such that $\nabla f(x) = a(x)$.

It is not hard to show that if both $f : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}$ are differentiable at $x \in \mathbb{R}^n$ then so is $c_1 f + c_2 g$, where c_1 and c_2 are numbers. Fortunately,

$$\nabla(c_1 f + c_2 g)(x) = c_1 \nabla f(x) + c_2 \nabla g(x).$$

In fact, the chain rule also applies: if $h : \mathbb{R} \to \mathbb{R}$, then

$$\nabla[h \circ f](x) = h'(f(x)) \nabla f(x). \tag{19}$$

Let $f : S \to \mathbb{R}$, where $S \subset \mathbb{R}^n$ is an open box. Let $e_j \in \mathbb{R}^n$ be the vector with 0s in every entry except for the jth, where the entry there is a 1. Then the jth **partial derivative** of f at the point $x \in S$ exists if for all $\epsilon > 0$ there is a $\delta > 0$ such that for any number t for which $x + te_j \in S$, $t < \delta$ implies

$$\left| \frac{f(x + te_j) - f(x)}{t} - a \right| < \epsilon \tag{20}$$

The number a, if it exists, is unique for each x and is the jth partial derivative. It defines the partial derivative function, $\frac{\partial f}{\partial x_j} : S \to \mathbb{R}$, a function defined by $\frac{\partial f(x)}{\partial x_j} = a$.

Similarly, if we replace every occurence of e_j in the definition of partial derivative by μ, where $\mu \in \mathbb{R}^n$ and restrict t to be positive, then we have the definition of "the **directional derivative** of f at x in the direction μ."

Theorem 9. *Consider $f : S \to \mathbb{R}$ where $S \subset \mathbb{R}^n$ is an open box. Then (i) if f is differentiable then it is continuous; (ii) if f is differentiable at x then $\partial f(x)/\partial x_j$ exist for all j and $\nabla f(x) = [\partial f(x)/\partial x_1, \ldots, \partial f(x)/\partial x_n]'$; (iii) if $\partial f(x)/\partial x_j$ exist for all j and are all continuous on x then $\nabla f(x)$ exists and is given by $\nabla f(x) = [\partial f(x)/\partial x_1, \ldots, \partial f(x)/\partial x_n]'$; (iv) if f is differentiable at x then the directional derivative of f exists for any vector μ and is equal to $\nabla f(x) \cdot \mu$.*

Exercise 26. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a function such that $f(0,0) = 0$, and for $(x, y) \neq 0$,

$$f(x, y) = \frac{xy}{\sqrt{x^2 + y^2}}.$$

Is f differentiable at $(0, 0)$?

Let $f : S \to \mathbb{R}$ where $S \subset \mathbb{R}^n$ is an open box. Suppose f is differentiable at $x \in S$, and suppose that each partial derivative function of f is differentiable at x. Denote the jth partial of $\partial f(x)/\partial x_i$ (also called the "(i, j)-cross partial") by $\partial^2 f(x)/\partial x_j \partial x_i$ if $j \neq i$ and $\partial^2 f(x)/\partial x_i^2$ if $j = i$. Then the **Hessian** of f at x is the matrix

$$Hf(x) = \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \cdots & \cdots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \cdots & \frac{\partial^2 f(x)}{\partial x_2^2} & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \cdots & \cdots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{bmatrix} \tag{21}$$

If every partial derivative of f is a continuous function, then we say that f is **continuously differentiable** or C^1.

If every (i, j)-cross partial of f is a continuous function then we say that f is C^2, and when f is C^2, it turns out that the Hessian is a symmetric matrix with

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i} \tag{22}$$

for all $i = 1, \ldots, n$ and $j = 1, \ldots, n$. This fact is called **Young's theorem**, and you will demonstrate it through an example below.

Let $f : S \to \mathbb{R}$, where $S \subset \mathbb{R}^n$ is an open box. Now let us treat x_j, $j \neq i$ as constants and define the function $g : S_i \to \mathbb{R}$ to be

$$g(x_i) \equiv f(x_i; x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n),$$

where the semicolon simply divides the free and fixed variables (i.e., "parameters"); alternatively, we may sometimes use " | " instead of the semi-colon to separate the free and fixed variables. Then you will be relieved to know that

$$\frac{\partial f}{\partial x_i} \equiv \frac{dg}{dx_i}. \tag{23}$$

So go ahead and use the chain rule, product rule, quotient rule, etc. that we described in the one variable case to calculate partial derivatives.

Exercise 27. Convince yourself that (22) and (23) hold.

Exercise 28. Let $f(x_1, x_2) = \log x_1(x_2)^2 + x_1 x_2$ and assume that f is C^2. Demonstrate Young's theorem for this function.