Week 6 Summary

In Section 1 we will discuss the last couple of items regarding differentiation. First, we talk about $C^k$ functions where $k = 0, 1, \ldots, \infty$. $C^k$ functions are the ones with graphs of varying degrees of “smoothness” with $C^\infty$ functions called smooth functions.

In Section 2, we will move on to a new topic of integration. We will define what it means for a function to be integrable. We will look at a couple of examples where we work directly with the definition of integrability of a function.

Topics

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Finishing up differentiation</td>
</tr>
<tr>
<td>1.1 $C^k$ functions and smooth functions</td>
</tr>
<tr>
<td>1.2 Newton’s Method</td>
</tr>
<tr>
<td>2 Integration</td>
</tr>
<tr>
<td>2.1 Definition</td>
</tr>
<tr>
<td>2.2 Examples</td>
</tr>
</tbody>
</table>

1 Finishing up differentiation

1.1 $C^k$ functions and smooth functions

Definition ($C^k$ functions)

Let $f$ be a function defined on an interval $I$. If $f$ is $k$-times differentiable with derivatives $f'(x), f''(x), \ldots, f^{(k)}(x)$ where $f^{(i)}(x)$ is continuous for all $i = 1, \ldots, k$, we say that $f(x)$ is $C^k$ on $I$.

If $f$ is $C^k$ for all $k$, then we say that $f$ is smooth on $I$ and write $f$ is $C^\infty$.

For example, $C^0$ functions are precisely those functions that are continuous. $C^1$ functions are those that have continuous first derivatives; $C^2$ are the functions with continuous second derivatives; and so on. In particular, note that if $k > k$ and $f$ is $C^k$, then $f$ is automatically $C^k$.

Intuitively, functions that are $C^k$ for high values of $k$ are the ones whose graphs are smoother. For example, let’s consider the following:

Example 1.1. Consider the function $f(x) = |x|$. Then, $f$ is $C^0$ but not $C^1$ on $\mathbb{R}$.

Proof. We already know that $|x|$ is continuous. Hence, $f$ is $C^0$. However, note that $f'(0)$ is not defined. Note that

$$f'(x) = \begin{cases} 1 & x > 0 \\ -1 & x < 0 \end{cases}$$

Hence, the left and the right limit of $f'(x)$ does not agree at 0. Therefore, $f$ is not differentiable at 0, and thus, is not $C^1$ on $\mathbb{R}$.

Example 1.2. Consider the function

$$f(x) = \begin{cases} x^2 & x \geq 0 \\ -x^2 & x < 0 \end{cases}$$

Then, $f$ is $C^1$ but not $C^2$. 

Proof. Since the right and the left limit of \( f(x) \) as \( x \to 0 \) are both 0, \( f(x) \) is clearly \( C^0 \). Now, we can compute the first derivative of \( f \):

\[
   f'(x) = \begin{cases} 
   2x & x \geq 0 \\
   -2x & x < 0 
   \end{cases}
\]

Note that \( f'(x) \) is defined and continuous on all of \( \mathbb{R} \). Hence, \( f \) is \( C^1 \). However, the second derivative of \( f \) is:

\[
   f''(x) = \begin{cases} 
   2 & x > 0 \\
   -2 & x < 0 
   \end{cases}
\]

Note that \( f''(0) \) does not exist since the left and the right limits do not agree. Therefore, \( f \) is \( C^1 \) but not \( C^2 \).

Recall the function \( f(x) = x \sin \left( \frac{1}{x} \right) \) was also continuous but not differentiable. In other words, \( f \) is \( C^1 \). By the same analysis, we can check that \( f_k(x) = x^{k+1} \sin \left( \frac{1}{x} \right) \) is \( C^k \) but not \( C^{k+1} \).

What about smooth functions? In fact, a lot of the functions that we deal with are smooth. Some examples of smooth functions are:

1. Polynomials
2. Exponential functions
3. Trigonometric functions
4. Logarithmic functions (on the positive reals)

All these functions are familiar to us and the fact that they are smooth is very helpful when working with them. However, sometimes we want to construct functions that behave the way we want. For example, we might want to require that a function be 0 for all \( x < 0 \). Let’s see how we can construct smooth functions that meet certain requirements.

Example 1.3. Let’s construct a smooth function \( f(x) \) such that \( f(x) = 0 \) for all \( x \leq 0 \) and \( f(x) > 0 \) for all \( x > 0 \).

Solution. We could try

\[
   f_1(x) = \begin{cases} 
   0 & x \leq 0 \\
   x & x > 0 
   \end{cases} \quad \text{or} \quad f_2(x) = \begin{cases} 
   0 & x \leq 0 \\
   x^2 & x > 0 
   \end{cases}
\]

Note that \( f_1 \) and \( f_2 \) both satisfy the condition that \( f(x) = 0 \) for all \( x \leq 0 \) and \( f(x) > 0 \) for \( x > 0 \). However, \( f_1 \) and \( f_2 \) are not smooth. In fact, \( f_1 \) is not \( C^1 \) and \( f_2 \) is not \( C^2 \).

Consider the following function \( f \):

\[
   f(x) = \begin{cases} 
   0 & x \leq 0 \\
   e^{-\frac{1}{x}} & x > 0 
   \end{cases}
\]

We claim that \( f(x) \) is a smooth function satisfying our conditions. First, since \( e^y \) is positive for all \( y \in \mathbb{R} \), \( f(x) > 0 \) for all \( x > 0 \). Hence, we only have left to show that \( f \) is smooth.

Since constant functions and exponential functions are smooth, we only need to check for continuity and differentiability at 0. First, we check that \( f(x) \) is continuous. To do that, we have to check that the left limit and the right limits of \( f(x) \) agree at 0. The left limit is simply \( \lim_{x \to 0^-} 0 = 0 \). The right limit is:

\[
   \lim_{x \to 0^+} e^{-\frac{1}{x}} = 0
\]
Hence, $f$ is $C^0$. Now, we take a look at $f'(x)$:

$$f'(x) = \begin{cases} 0 & x \leq 0 \\ \frac{1}{x^2} e^{-\frac{x}{2}} & x > 0 \end{cases}$$

Again, we have to check that the right limit of $f'(x) = 0$. In other words, we have to check that

$$\lim_{x \to 0} \frac{e^{-\frac{x}{2}}}{x^2} = 0$$

However, this is true since $e^{-\frac{x}{2}}$ is $o(x^n)$ for all $n$. Therefore, $f'(x)$ is continuous and thus, $f$ is $C^1$.

Since $e^{-\frac{x}{2}}$ is $o(x^n)$ for all $n$, by similar argument, we can show that $f^{(n)}(x)$ is continuous for all $n$. Therefore, we conclude that $f(x)$ is indeed smooth.

Now, using above $f(x)$ as a building block, we can construct other smooth functions that meet other conditions.

**Example 1.4.** Construct a smooth function $g(x)$ such that

1. $g(x) = 0$ for all $x \leq 0$
2. $g(x) = 1$ for all $x \geq 1$
3. $0 < g(x) < 1$ for all $0 < x < 1$

**Solution.** Consider the function

$$g(x) = \frac{f(x)}{f(x) + f(1-x)}$$

where $f(x)$ is the function from Example 1.3. We claim that $g(x)$ satisfies all the conditions listed above.

First, note that $f(x) + f(1-x) \neq 0$ for all $x$: Suppose $f(x) + f(1-x) = 0$. Since $f$ is always nonnegative, this means that $f(x) = 0$ and $f(1-x) = 0$. In particular, $x \leq 0$ and $1-x \leq 0$. Therefore, we have $1 \leq x \leq 0$ which is a contradiction. Hence, $f(x) + f(1-x) \neq 0$ for all $x$.

Now, $g(x)$ is a quotient of smooth functions whose denominator is nonzero. Hence, $g(x)$ is also smooth. Let’s check that $g(x)$ satisfies the conditions above.

1. Let $x \leq 0$. Then,

$$g(x) = \frac{f(x)}{f(x) + f(1-x)} = \frac{0}{0 + f(1-x)} = 0$$

2. Let $x \geq 1$. Then,

$$g(x) = \frac{f(x)}{f(x) + f(1-x)} = \frac{f(x)}{f(x) + 0} = 1$$

3. Let $0 < x < 1$. Note that $f(x) \geq 0$ for all $x$ and $f(x) + f(1-x) > 0$ for all $x$. Hence,

$$f(x) \geq f(x) + f(1-x)$$

with equality if and only if $f(1-x) = 0$ i.e. $x \geq 1$. In particular, when $x < 1$, $f(x) < f(x) + f(1-x)$. Therefore, $g(x) < 1$ when $x < 1$. Moreover, since $f(x) > 0$ and $f(x) + f(1-x) > 0$ for all $x > 0$, we know that $g(x) > 0$.

Thus, for all $0 < x < 1$, we have that $0 < g(x) < 1$.

Therefore, $g(x)$ is a smooth function satisfying all the conditions above.
1.2 Newton’s Method

Until now, we have covered several applications of derivatives such as the Mean Value Theorem. However, these “applications” might not have seemed too useful or really applicable to real world situations (except maybe the highway patrol could have used the MVT to give me a citation for reckless driving).

Newton’s method will have more practical application. More precisely, Newton’s method helps us approximate solutions to equations, which is especially helpful when we cannot explicitly solve the equation.

So how does it work? Suppose we have a function \( f(x) \) and we would like to find a solution for \( f(x) = 0 \). First, we make an initial guess, say \( x_0 \). If we are lucky, we will have \( f(x_0) = 0 \) and we will have found our solution. If not, we consider

\[
x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}
\]

Note that \( x_1 \) is precisely the \( x \)-intercept of the line \( y = f(x_0) + f'(x_0)(x - x_0) \). However, this is precisely the equation of the tangent line to \( f(x) \) at \( f(x_0) \). Under certain conditions that we will see later, \( f(x_1) \) is guaranteed to be closer to 0 than \( f(x_0) \). In other words, \( x_1 \) is a better approximation of the solution to \( f(x) = 0 \) than \( x_0 \).

Continuing in the same way, we define for each \( n \)

\[
x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}
\]

Again, under certain conditions, we know that \( x_n \) will converge to a limit \( x^* \) such that \( f(x^*) = 0 \).

Why is this useful? For example, suppose we want to compute \( \sqrt{2} \). Recall that \( \sqrt{2} \) is defined to be the least upper bound of the set \( \{x \in \mathbb{R} \mid x^2 \leq 2\} \). Finding the decimal representation of \( \sqrt{2} \) using this abstract definition is just impossible. Instead, we can just try to find a real number \( x \) such that \( x^2 = 2? \) That will work but we quickly see that immediately coming up with the actual solution to this equation is very hard. We can use Newton’s method to approximate the solution, which we know will be \( \sqrt{2} \).

In fact, this was one of the interview questions for Bloomberg’s software engineer position this year. So if nothing else, Newton’s method can help you with these interview questions!

Example 1.5. Approximate \( \sqrt{e} \) using Newton’s method (assuming we know the value of \( e \)).

Solution. First, we need to put the question in the right framework to be able to use Newton’s method. Our function \( f(x) \) has to be a function such that \( f(\sqrt{e}) = 0 \). There are lots of such functions but let’s choose the simplest such function i.e. \( f(x) = x^2 - e \).

Note that \( f'(x) = 2x \). Hence, we already have \( x_n \), we compute

\[
x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{x_n^2 - e}{2x_n} = \frac{x_n^2 + e}{2x_n}
\]

Now, let’s start by making our initial guess, \( x_0 \). Let’s start with a pretty terrible guess: \( x_0 = 1 \). Clearly, \( 1^2 = 1 < e \) and we definitely could have chosen a better initial guess. However, we will see that even with this terrible guess, Newton’s method will give us a pretty good approximation of \( \sqrt{e} \) quite quickly.

Using above computation, we get:

\[
x_1 = \frac{1 + e}{2} \approx 1.86 \quad ; \quad x_2 = \left( \frac{1 + e}{2} \right)^2 + 2 \right) / (1 + e) \approx 1.66 \quad ; \quad x_3 = \cdots \approx 1.64876
\]

The true value of \( \sqrt{e} = 1.64872 \ldots \) so only after 3 iterations of Newton’s method starting with a bad initial guess, we got a value that is correct up to the fourth decimal place!.

Even though Newton’s method worked quite brilliantly in above example, it can sometimes fail miserably as well.
Example 1.6. Let’s find the root of \( f(x) = \sqrt[3]{x} \) using Newton’s method.

Solution. Let’s ignore for a moment that the solution to \( f(x) = 0 \) is quite clearly \( x = 0 \). We again start with a bad guess \( x_0 = 1 \). We first find the recursive formula for \( x_n \):

\[
f'(x) = \frac{1}{3} x^{-2/3} \Rightarrow x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - 3 \frac{x_n^{-1/3}}{x_n^{-2/3}} = -2x_n
\]

This means that if we set \( x_0 = 1 \) we get

\[
x_1 = -2 \quad ; \quad x_2 = 4 \quad ; \quad x_3 = -8 \cdots
\]

In this example, instead of getting closer and closer to the solution, which is \( x = 0 \), Newton’s method gave us a sequence of \( x_n \)’s which got farther and farther away. So Newton’s method fails quite spectacularly in this example.

Why did Newton’s method work in the first example and fail in the second? More importantly, can we tell when Newton’s method is going to work? Fortunately, the following theorem will tell us when Newton’s method will work.

**Theorem 1.1.** Let \( I \) be an interval and \( f \) a \( C^2 \) function on \( I \). Suppose that for every \( x \in I \), we have \( |f''(x)| < M \) and \( |f'(x)| > \frac{1}{K} \) for some \( M \) and \( K \). Let \( x_0 \in I \), and define the sequence \( \{x_n\} \) by

\[
x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}
\]

If \( x_n \in I \) for each \( N \) and \( |f(x_0)| < \frac{r}{K^2 M} \), then for all \( n \) we have

\[
|f(x_n)| \leq \frac{r^{2^n}}{K^2 M}
\]

2 Integration

Now, we move on from differentiation to integration. You might have learned in high school that integral of \( f \) over the interval \([a, b]\) is the “area under the graph of \( f \) over \([a, b]\)”. In this section, we will make that idea precise using our new best friend \( \varepsilon > 0 \).

2.1 Definition (Partition of an interval)

Let \( I = [a, b] \) be an interval. A partition of \( I \) is a set \( P = \{t_0, t_1, \ldots, t_n\} \) such that

\[
a = t_0 < t_2 < t_3 < \cdots < t_{n-1} < t_n = b
\]

Given any interval \([a, b]\), one useful partition of \([a, b]\) that we will use repeatedly is the **uniform partition**, which partitions \([a, b]\) into \( n \) subintervals of equal length. In particular, we get

\[
P_n = \{t_0, \ldots, t_n\} \quad \text{where} \quad t_i = a + \frac{b-a}{n}
\]

Once we have chosen a partition of the interval, we can define the **Riemann upper/lower sums**.
Definition (Riemann upper/lower sum)

Let $\mathcal{P}$ be a partition of $[a, b]$. Then, for a function $f$ which is bounded on $[a, b]$ we define:

$$U_{\mathcal{P}}(f) := \sum_{i=1}^{n} \left( \sup_{x \in (t_{i-1}, t_i)} f(x) \right) (t_i - t_{i-1})$$

$$L_{\mathcal{P}}(f) := \sum_{i=1}^{n} \left( \inf_{x \in (t_{i-1}, t_i)} f(x) \right) (t_i - t_{i-1})$$

(Here, sup means “least upper bound” and inf means “greatest lower bound”). $U_{\mathcal{P}}(f)$ is called the upper Riemann sum of $f$ with respect to $\mathcal{P}$; and $L_{\mathcal{P}}(f)$ is called the lower Riemann sum of $f$ with respect to $\mathcal{P}$.

Looking at Figure 2.1, $L_{\mathcal{P}}$ is the sum of the area of the blue rectangles (shaded blue) and $U_{\mathcal{P}}$ is the sum of the area of the red rectangles (shaded blue + grey). Clearly, $U_{\mathcal{P}} \geq L_{\mathcal{P}}$ for all partitions $\mathcal{P}$. Intuitively, we will say that a function $f$ is integrable if by partitioning our interval “finer”, the two values of $U_{\mathcal{P}}$ and $L_{\mathcal{P}}$ converge to the same limit.

![Figure 1: The upper and lower Riemann sums](image)

Definition (Integrability of $f$)

Let $f$ be a bounded function on an interval $I = [a, b]$. We say that $f$ is integrable on $I$ if for all $\varepsilon > 0$, we can find a partition $\mathcal{P}$ of $I$ such that $U_{\mathcal{P}} - L_{\mathcal{P}} < \varepsilon$.

If $f$ is integrable on $I$, we write

$$\int_{a}^{b} f(x)dx = \lim_{n \to \infty} U_{\mathcal{P}_n}(f) = \inf \{ U_{\mathcal{P}} \mid \mathcal{P} \text{ is a partition of } I \}$$

where $\mathcal{P}_n$ is the uniform partition of $I$ into $n$ subintervals.

You might note that this definition is slightly different from the one given in the course notes for Ma 1a but in fact, the two definitions are equivalent, and virtually identical. Also, this definition tells us that when the function is integrable, the integral of the function is indeed the “area under the curve” which is the graph of $f$.

2.2 Examples

Working with the definition of integral is quite a pain. We will develop tools that will allow us work with integrals without directly working with $\varepsilon$. But first, let’s see how we can work with the definition:

Example 2.1. Let $f(x) = c$. Show that $f$ is integrable on any interval $[a, b]$ and that

$$\int_{a}^{b} f(x)dx = c(b - a)$$
Proof. First, let’s see how the upper and lower Riemann sums of a fixed partition $P$ looks like. Let’s look at the uniform partition $P_n$, i.e.

$$P_n := \{a, a + \frac{b-a}{n}, a + 2\frac{b-a}{n}, a + 3\frac{b-a}{n}, \ldots, a + n\frac{b-a}{n} = b\}$$

Now, let’s first compute sup and inf of $f(x)$ over each subinterval $(t_i, t_{i+1})$. Note that since $f(x) = c$ is a constant function, we have

$$\sup_{x \in (t_i, t_{i+1})} f(x) = \inf_{x \in (t_i, t_{i+1})} f(x) = c$$

for all $i$. Hence, we get:

$$U_{P_n}(f) = \sum_{i=1}^{n} \left( \sup_{x \in (t_{i-1}, t_i)} f(x) \right) (t_i - t_{i-1}) = \sum_{i=1}^{n} c \cdot \frac{b-a}{n} = c(b-a)$$

Similarly, $L_{P_n}(f) = c(b-a)$. Hence, for all $n$, we have $U_{P_n}(f) = L_{P_n}(f) = c(b-a)$. Therefore, we conclude that

$$\int_{a}^{b} f(x)dx = c(b-a)$$

That was an easy example. Now, let’s look at a slightly harder example.

**Example 2.2.** Let $f(x) = x^p$ where $p \in \mathbb{N}$. Show that for any $b > 0$, $f(x)$ is integrable on the interval $[0, b]$ and that

$$\int_{0}^{b} f(x)dx = \frac{b^{p+1}}{p+1}$$

Proof. Again, we begin by taking a look at the uniform partition $P_n$ of the interval $[0, b]$. Since the interval is $[0, b]$, we have

$$P_n = \{0, \frac{b}{n}, \frac{2b}{n}, \ldots, \frac{nb}{n} = b\}$$

Now, since $f(x) = x^p$ is an increasing function on $[0, \infty)$, given any interval $(t_{i-1}, t_i)$, we know that

$$\sup_{(t_{i-1}, t_i)} f(x) = t_i^p \quad \text{and} \quad \inf_{(t_{i-1}, t_i)} f(x) = t_{i-1}^p$$

Hence, we obtain

$$U_{P_n}(f) = \sum_{i=1}^{n} \left( \sup_{(t_{i-1}, t_i)} x^p \right) (t_i - t_{i-1}) = \sum_{i=1}^{n} \left( \frac{ib}{n} \right)^p \cdot \frac{b}{n} = \sum_{i=1}^{n} \frac{b^p}{n^p+1} \cdot \frac{b^{p+1}}{n^{p+1}} = \frac{b^{p+1}}{n^{p+1}} \cdot \sum_{i=1}^{n} \frac{b^p}{i^p}$$

$$\int_{0}^{b} f(x)dx = \frac{b^{p+1}}{p+1}$$
Similarly, we can compute $L_{p_n}$

\[
L_{p_n}(f) = \sum_{i=1}^{n} \left( \inf_{(t_{i-1}, t_i)} x^p \right) (t_i - t_{i-1})
\]

(5)

\[
= \sum_{i=1}^{n} \left( \frac{(i-1) b}{n} \right)^p \frac{b}{n}
\]

(6)

\[
= \sum_{i=1}^{n} (i - 1)^p \cdot \left( \frac{b^{p+1}}{n^{p+1}} \right)
\]

(7)

\[
= \frac{b^{p+1}}{n^{p+1}} \sum_{i=1}^{n} (i - 1)^p
\]

(8)

Now, note that

\[
U_{p_n} - L_{p_n} = \left( \frac{b^{p+1}}{n^{p+1}} \cdot \sum_{i=1}^{n} i^p \right) - \left( \frac{b^{p+1}}{n^{p+1}} \cdot \sum_{i=1}^{n} (i - 1)^p \right)
\]

(9)

\[
= \frac{b^{p+1}}{n^{p+1}} \cdot \sum_{i=1}^{n} (i^p - (i - 1)^p)
\]

(10)

\[
= \frac{b^{p+1}}{n^{p+1}} \cdot n^p
\]

(11)

\[
= \frac{b^{p+1}}{n}
\]

(12)

In particular, note that

\[
\lim_{n \to \infty} (U_{p_n} - L_{p_n}) = \lim_{n \to \infty} \frac{b^{p+1}}{n} = 0
\]

This is great news! By making $n$ large enough, we can make $U_{p_n} - L_{p_n}$ very small.

Let $\varepsilon > 0$. Choose $N$ large enough so that $\frac{b^{p+1}}{N} < \varepsilon$. Then,

\[
U_{p_N} - L_{p_N} = \frac{b^{p+1}}{N} < \varepsilon
\]

Therefore, $f(x)$ is integrable on the interval $[0, b]$. Moreover, we can compute the integral of $f(x)$ over $[0, b]$:

\[
\int_0^b f(x) \, dx = \lim_{n \to \infty} U_{p_n}(f)
\]

(13)

\[
= \lim_{n \to \infty} \left( \frac{b^{p+1}}{n^{p+1}} \cdot \sum_{i=1}^{n} i^p \right)
\]

(14)

\[
= \frac{b^{p+1}}{n^{p+1}} \cdot \frac{1}{p+1} \sum_{i=1}^{n} i^p
\]

(15)

\[
= \frac{b^{p+1}}{p+1}
\]

(16)

where the limit in line (15) was calculated using one of your past homework problems for Ma 1a: Recall that we defined

\[
S_k(n) = \sum_{i=1}^{n} i^k
\]
and we showed that $S_k(n)$ is a polynomial in $n$ whose leading term is $\frac{n^{k+1}}{k+1}$. In particular,

\[
\lim_{n \to \infty} \frac{1}{n^{p+1}} \sum_{i=1}^{n} i^p = \lim_{n \to \infty} \frac{1}{n^{p+1}} S_p(n) = \lim_{n \to \infty} \frac{n^{p+1} + \text{(lower order terms)}}{n^{p+1}} = 1 \quad \text{(18)}
\]

\[
= \frac{1}{p+1} \quad \text{(19)}
\]

\[= \frac{1}{p+1} \quad \text{(20)}
\]