MAT137 - Term 2, Week 10

- Problem Set 9 is due tomorrow, 17 March, at 3pm.
- Today we will:
  - Remind ourselves a bit about our discussion at the end of last class.
  - Talk about power series (more formally this time).
  - Talk about Taylor polynomials.
  - Introduce the idea of a Taylor series (if there's time).

Discussion from the end of last class

At the end of last class, we spent some time discussing the following problem.

Suppose you're a 17^{th} century mathematician (who doesn't already know that the answer to this question is e^x), and you're trying to find a function f that satisfies the following two conditions:

$$f' = f \quad \text{and} \quad f(0) = 1.$$

Immediately we realized that from these conditions we can conclude:

$$f'(0) = f(0) = 1, \quad \text{and in general} \quad f^{(n)}(0) = 1, \quad \text{for all } n \in \mathbb{N}$$

Discussion from the end of last class

Since we're in the 17th century and fancy functions like sines, cosines, and exponentials haven't been invented yet, we're trying to find a polynomial that works.

Quickly we were able to conclude that if p_n is a polynomial with degree n, in order for it to satisfy the first $n+1$ of these conditions:

$$p_n(0) = p_n'(0) = p_n''(0) = \cdots = p_n^{(n-1)}(0) = p_n^{(n)}(0) = 1.$$

we must have:

$$p_n(x) = 1 + x + \frac{1}{2}x^2 + \frac{1}{3 \cdot 2}x^3 + \cdots + \frac{1}{n!}x^n.$$

But, we will always have $p_n^{(n+1)}(0) = 0 \neq 1$, so this function doesn't work.

Discussion from the end of last class

Then we said to ourselves, "Well, what if the polynomial *just keeps going*...?"

That led us to try to define a function like this:

$$f(x) = 1 + x + \frac{1}{2}x^2 + \frac{1}{3 \cdot 2}x^3 + \cdots + \frac{1}{n!}x^n + \cdots .$$

This doesn't make sense yet, exactly, but we decided to work with it anyway. We asked ourselves what happens if we could differentiate this function "term by term" like we do with an ordinary polynomial.

We did this and concluded that if that makes sense, then $f'(x) = f(x)$.

So this seems to work! Today we'll try to figure out whether this even makes sense.

Power series

The idea with power series is to use a series to define a function.

We've seen two examples of this so far. One is the idea we were just talking about, where we defined

$$f(x) = 1 + x + \frac{1}{2}x^2 + \frac{1}{3 \cdot 2}x^3 + \cdots + \frac{1}{n!}x^n + \cdots = \sum_{n=0}^{\infty} \frac{1}{n!}x^n.$$

The other example we've seen is from the geometric series formula:

$$\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n \quad \text{as long as} \quad |x| < 1.$$

Power series

The first question you should ask is: Why would we do this?

The discussion from last class provides part of the answer.

- If these functions end up working like polynomials, then they might be easy to differentiate (and therefore also easy to integrate).
- We can design such a function to have specific properties which might be useful, like we did when we designed a function to satisfy the differential equation $f'(x) = f(x)$.

The geometric series example tells us that we have to be careful though. A series involving x might only definite a function on a certain domain.

So, we'll proceed cautiously.

Definition of power series

Definition

Let $\{a_n\}$ be a sequence of real numbers, and let $a \in \mathbb{R}$.

A series of the form $\sum a_n (x-a)^n$ is called a power series.

More specifically, it's called a power series in $(x-a)$ or a power series centred at a (for reasons that will become clear later).

The specific case of power series cented at $a = 0$ will be what we study most carefully. These power series look like this:

$$\sum a_n x^n.$$

Convergence of power series

If you take a power series and substitute in a particular value of x, you get a "regular" series with no more variables.

Definition

- Let $c \in \mathbb{R}$. A power series $\sum a_n(x-a)^n$ is said to converge at c if $\sum a_n(c-a)^n$ converges.
- Let S be a set of real numbers. A power series $\sum a_n(x-a)^n$ is said to converge on S if it converges at c for every $c \in S$.

Power series as functions

As we said, the idea here is to define functions with power series. So you should think about defining a function f via something like:

$$f(x) = \sum_{n=0}^{\infty} a_n (x-a)^n.$$

As our geometric series example shows us, we have to be careful with the domain of this function. It's very possible for the series on the right to converge for some values of x and not for others.

We can say the following, not very useful fact:

The domain of a power series $\sum a_n (x-a)^n$ is the set of all $c \in \mathbb{R}$ such that $\sum a_n (c-a)^n$ converges.

Power series as functions

Let's try to say something more useful.

We'll state this result for power series centred at 0 first.

Theorem

- *If the power series* $\sum a_n x^n$ *converges at some* $c \neq 0$, *then it converges absolutely for all* x *such that* $|x| < |c|$.
- *If* $\sum a_n x^n$ *diverges at some* d, *then it diverges for all* x *such that* $|x| > |d|$.

Power series as functions

We'll give the proof of the first part. Suppose $\sum a_n c^n$ converges.

Then by the NCT, we must have that $\lim_{n\to\infty} a_n c^n = 0$.

By definition of sequence convergence, there must be an N such that $|a_n c^n| < 1$ for all $n > N$.

Now let x be any real number such that $|x| < |c|$. We'd like to show that $\sum a_n x^n$ converges absolutely. We'll do this with the BCT.

$$|a_n x^n| = |a_n c^n| \left|\frac{x^n}{c^n}\right| < \left|\frac{x^n}{c^n}\right| = \left|\frac{x}{c}\right|^n \quad \text{for all} \quad n > N.$$

$\sum \left|\frac{x}{c}\right|$ is a convergent geometric series since $|x| < |c|$, so we're done by the BCT.

As a result of the previous theorem, the set of points on which a power series $\sum a_n x^n$ converges can only have three forms:

Case 1: The power series converges only when $x = 0$.

Obviously every series $\sum a_n x^n$ converges when $x = 0$.

An example where it converges nowhere else might be $\sum n!\, x^n$.
For any $x \neq 0$, we have that

$$\lim_{n\to\infty} n!\, x^n \quad \text{does not exist by the Big Theorem.}$$

Case 2: The power series converges *absolutely* at all real numbers.

We know an example of this already: $\sum \frac{x^n}{n!}$.

We can prove this with the ratio test:

$$\left| \frac{x^{n+1}}{(n+1)!} \cdot \frac{n!}{x^n} \right| = \left| \frac{x}{n+1} \right| = \frac{1}{n+1}|x| \to 0 \quad \text{for all } x \in \mathbb{R}.$$

Intervals of convergence

This leaves the more interesting case:

Case 3: There is a positive real number R such that the power series $\sum a_n x^n$ converges absolutely when $|x| < R$, and diverges when $|x| > R$.

We will soon see that when $|x| = R$ in this case, many things can happen. For now let's ignore that part.

Radius of convergence

Definition

Associated to every power series $\sum a_n x^n$ is a radius of convergence.

1. In Case 1 from above, we say the radius of convergence is 0.
2. In Case 2 from above, we say the radius of convergence is ∞.
3. In Case 3 from above, we say the radius of convergence is R.

What about the endpoints?

In Case 3, anything can happen at the endpoints:

1. $\sum x^n$ converges on $(-1, 1)$.
2. $\sum \frac{(-1)^n}{n} x^n$ converges on $(-1, 1]$.
3. $\sum \frac{1}{n} x^n$ converges on $[-1, 1)$.
4. $\sum \frac{1}{n^2} x^n$ converges on $[-1, 1]$.

What about when $a \neq 0$?

If your power series is not centred at zero, so it looks like $\sum a_n (x-a)^n$ for some $a \neq 0$, all the same stuff is true if you replace x with $x - a$.

So a series like this can:

1. Converge only at $x = a$.
2. Converge absolutely at all real numbers.
3. Converge absolutely when $|x - a| < R$ for some R, and diverge when $|x - a| > R$. (And again, any combination of things can happen when $|x - a| = R$.)

This shows that the set on which a power series converges must be an interval centred at a. We call it the interval of convergence.

This is also why we call it a power series centred at a.

Examples

Let's try to find the intervals of convergence of some power series.

The usual method here is to use the ratio test to get the radius of convergence, then analyze the two endpoints separately.

Example: $\sum \frac{(-1)^n}{\sqrt{n}} x^n$. (Interval of convergence: $(-1, 1]$.)

Example: $\sum \frac{3^n}{(3n)!} x^n$. (Interval of convergence: $\mathbb{R}$.)

Example: $\sum \frac{1}{n\,7^n} x^n$. (Interval of convergence: $\left[-\frac{1}{7}, \frac{1}{7}\right]$.)

Example: $\sum \frac{n^7}{e^n} (x-4)^n$. (Interval of convergence: $(4-e, 4+e)$.)

The idea we explored with our "17^{th} century example" was that power series are sort of like "infinite degree polynomials".

Our hope was that we could differentiate (and integrate) them exactly like they're polynomials.

For example, wouldn't it be nice if:

$$\frac{d}{dx}\left(\sum_{n=1}^{\infty} a_n\, x^n\right) = \sum_{n=1}^{\infty} \frac{d}{dx}\left(a_n\, x^n\right) = \sum_{n=1}^{\infty} n\, a_n\, x^{n-1}?$$

Then we could differentiate and integrate these functions very easily.

Differentiation and integration

The moral of the story is that most of the time we can do this:

<u>Inside</u> the interval of convergence (ie. when $|x| < R$), you can treat a power series like a polynomial.

To state this more precisely, there are two parts. If a power series converges absolutely at x, then

1. If you differentiate a power series term by term, then the resulting series converges.

2. The resulting series is actually the derivative of the series you started with.

Differentiation and integration

Here's the first part. Again, we'll only state this for power series centred at 0, but it applies to all power series with the appropriate shift.

Theorem

Let $\sum a_n x^n$ be a power series with radius of convergence R (which could be ∞).

1. *The power series $\sum \frac{d}{dx}(a_n x^n) = \sum n\, a_n x^{n-1}$ also converges when $|x| < R$.*

2. *The power series*

$$\text{“}\sum \left(\int a_n x^n \, dx \right)\text{”} = \sum \frac{a_n}{n+1} x^{n+1}$$

also converges when $|x| < R$.

Differentiation

Here's the second part, which is very surprising, and very useful.

We'll state it for derivatives first.

Theorem

Let $\sum a_n x^n$ be a power series with radius of convergence R, and define:

$$f(x) = \sum_{n=0}^{\infty} a_n x^n \quad \text{for all } x \in (-R, R).$$

Then f is differentiable on $(-R, R)$, and

$$f'(x) = \sum_{n=0}^{\infty} \frac{d}{dx}(a_n x^n) = \sum_{n=1}^{\infty} n\, a_n x^{n-1} \quad \text{for all } x \in (-R, R).$$

In fact, this immediately implies that f is differentiable infinitely many times on $(-R, R)$.

Integration

Theorem

Let $\sum a_n x^n$ *be a power series with radius of convergence* R*, and define:*

$$f(x) = \sum_{n=0}^{\infty} a_n x^n \quad \text{for all } x \in (-R, R).$$

Then f *is integrable on* $(-R, R)$*, and the function*

$$F(x) = \sum_{n=0}^{\infty} \frac{a_n}{n+1} x^{n+1} \quad \text{is an antiderivative of } f \text{ on } (-R, R).$$

You can do even more stuff

You can even multiply power series, though it's pretty tedious.

Suppose we define two functions with power series (centred at the same point):

$$f(x) = \sum_{n=0}^{\infty} a_n x^n \quad \text{and} \quad g(x) = \sum_{n=0}^{\infty} b_n x^n$$

Also suppose that x is some point <u>inside</u> both of their intervals of convergence. Then:

$$f(x)g(x) = \left[a_0 + a_1 x + a_2 x^2 + \dots\right]\left[b_0 + b_1 x + b_2 x^2 + \dots\right]$$

You can expand and collect terms:

$$= a_0 b_0 + (a_1 b_0 + a_0 b_1)x + (a_2 b_0 + a_1 b_1 + a_0 b_2)x^2 + \cdots$$

and this power series will also converge absolutely.

Absolutely convergent power series are great

This is the moral of the story:

Inside the interval of convergence (ie. when $|x| < R$), where the power series converges absolutely, you can treat a power series like a polynomial.

This is the main reason we care about absolute convergence so much.

Taylor polynomials

Now that we know about power series, we want to use them to help us understand functions better and understand series better.

All of this will come via our discussion of Taylor series.

The idea throughout this topic will be to take messy functions we don't know much about, and approximate them with nice functions (ie. polynomials).

We already did this with the function e^x with our example at the end of last class.

Taylor polynomials

Example: Let f be a function that is differentiable everywhere.

Suppose we want to approximate f near $x = a$. That is, come up with a function g whose values are "close to" the values of f for points near $x = a$.

One easy way to do this is to let g be the tangent line to the graph of f:

$$g(x) = f'(a)(x - a) + f(a).$$

We can see that the approximation is perfect at a:

$$g(a) = f'(a)(a - a) + f(a) = f(a).$$

Taylor polynomials

$$g(x) = f'(a)(x - a) + f(a).$$

However, there's no reason to believe that $g(x) = f(x)$ for any other points, so our approximation will likely have some error. We give that error a name:

$$R(x) = f(x) - g(x).$$

There's no error at a, so $R(a) = 0$, but for $x \neq a$, we will probably have $R(x) \neq 0$.

We want our approximation to get better the closer x is to a. So we may want to require:

$$\lim_{x \to a} R(x) = 0.$$

ie. We want the error to get smaller the closer x is to a.

Even more than that though, we want the error to approach zero fast. ie. We want the approximation to get *much* better the closer we are to a.

Definition

Let f and g be two functions such that $\lim_{x \to a} f(x) = \lim_{x \to a} g(x) = 0$ for some real number a.

We say g approaches zero faster than f as $x \to a$ if

$$\lim_{x \to a} \frac{g(x)}{f(x)} = 0.$$

This definition gives us a way to compare two functions, and see which one is going to zero faster than the other.

Taylor polynomials

Now we can say one function goes to zero faster than another.

We want to make sure our error $R(x)$ approaches zero very quickly at a, so we need something to measure it against; some "standard" family of functions that can serve as a measuring stick for our error functions.

Luckily, we have a very convenient family of functions around that can do this for us:

$$(x-a), (x-a)^2, (x-a)^3, \dots, (x-a)^k, \dots$$

Each of these functions approaches zero at a, and each function on the list approachees zero faster than all the functions before it on the list:

$$\text{For any integers } m < n, \text{ we have } \lim_{x\to a} \frac{(x-a)^n}{(x-a)^m} = \lim_{x\to a}(x-a)^{n-m} = 0.$$

Taylor polynomials

These functions $(x-a)^k$ will serve as our "standard" family of functions.

Essentially we'll use them to say something like "$R(x)$ approaches zero as $x \to a$ at least 'n-quickly' if $\lim_{x\to a} \frac{R(x)}{(x-a)^n} = 0$."

Here's the more precise version, in the context of approximations.

Definition

Let f and g be two functions such that $\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = 0$ for some real number a.

We say <u>g is a good approximation of order n for f at a</u> if

$$\lim_{x\to a} \frac{f(x)-g(x)}{(x-a)^n} = 0.$$

Taylor polynomials

So a good approximation of order n for f at a is another function g such that the error in the approximation:

$$R(x) = f(x) - g(x),$$

goes to zero faster than the n^{th} function in our list of standard functions.

So what does this all have to do with polynomials? We're getting there.

Theorem

Let a be a real number, n a positive integer, and let f and g be two functions whose first n derivatives agree at a:

$$f(a) = g(a), f'(a) = g'(a), \cdots, f^{(n-1)}(a) = g^{(n-1)}(a), f^{(n)}(a) = g^{(n)}(a).$$

Then g is a good approximation of order n for f at a.

Taylor polynomials

The previous theorem follows easily from l'Hôpital's rule:

Suppose the first n derivatives of f and g agree at a. Then in particular we know that $f(x) - g(x)$ is continuous, and

$$\lim_{x \to a} f(x) - g(x) = 0.$$

Therefore, by l'Hôpital's rule, we have:

$$\lim_{x \to a} \frac{f(x) - g(x)}{(x-a)^n} = \lim_{x \to a} \frac{f'(x) - g'(x)}{n(x-a)^{n-1}}.$$

Since by assumption $f'(x) - g'(x)$ is differentiable, it is continuous, and we can apply l'Hôpital's rule again to this limit on the right. Doing this n times, we eventually arrive at:

$$\lim_{x \to a} \frac{f(x) - g(x)}{(x-a)^n} = \lim_{x \to a} \frac{f'(x) - g'(x)}{n(x-a)^{n-1}} = \cdots = \lim_{x \to a} \frac{f^{(n)}(x) - g^{(n)}(x)}{n!} = 0.$$

Taylor polynomials

So finally, the last theorem gives us a simple condition we can check to see whether a given function is a good approximation for another function.

We now apply this framework to polynomials.

Definition

Let f be a function that has all of its derivatives, let a be a real number, and n a positive integer.

The n^{th} Taylor polynomial of f at a is the unique polynomial P_n of smallest degree such that

$$P_n(a) = f(a), P_n'(a) = f(a), \dots, P_n^{(n-1)}(a) = f^{(n-1)}(a), P_n^{(n)}(a) = f^{(n)}(a)$$

Taylor polynomials

Note that the definition does not require that P_n is a degree n polynomial. Its degree can be less.

It is a routine exercise to derive the following formula (which is just a generalization of what we did in the differential equation example last class:

Theorem

Let f be a function that has all of its derivatives, let a be a real number, and n a positive integer.

The n^{th} Taylor polynomial of f at a has the following form:

$$P_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k$$