14 Taylor polynomials and Taylor's theorem

14.1 Definition of the Taylor polynomial

Suppose we know a lot about a function f at a point a (the function value, in some exact form, the value of the derivative, et cetera), but don't have such easy access to values away from a. Can we use the information we have at a to say *something* about the function away from a?

Examples of this type of situation include:

- $f(x) = \sin x$ at 0 (we know *everything* about the function at 0, in a quite exact way, but very little about it away from 0)
- $f(x) = \sin x$ at π, or $-\pi$, or $3\pi/2$, ...
- $f(x) = \log x$ at 1
- $f(x) = \sqrt{x^2 + 9}$ at 4, or 0, or -4.

There's an obvious, but next-to-useless, way to approximate f near a, using data at a — just use the constant function $f(a)$. A less obvious, and much more useful, way, is to use the linearization of f at a to approximate f near a, that is, to use the function

$$f(a) + f'(a)(x - a),$$

which has the property that it agrees with f at a, and also agrees with f' at a, so its graph agrees with the graph of f at a, and is also "traveling in the same direction" as the graph of f at a.

We can push this further: the function

$$f(a) + f'(a)(x - a) + \frac{f''(a)}{2}(x - a)^2$$

is easily seen to agree with each of f, f' and f'' at a, and more generally the function

$$f(a) + f'(a)(x - a) + \frac{f''(a)}{2}(x - a)^2 + \ldots + \frac{f^{(n)}(a)}{n!}(x - a)^n$$

is easily seen to agree with each of $f, f', f'', \ldots, f^{(n)}$ at a.[228]

This example leads to the definition of the *Taylor polynomial*.

[228] One way to prove this formally is to prove by induction on k that for $n \geq k \geq 0$, the kth derivative of

$$f(a) + f'(a)(x - a) + \frac{f''(a)}{2}(x - a)^2 + \ldots + \frac{f^{(n)}(a)}{n!}(x - a)^n$$

is

$$\sum_{j=k}^{n} \frac{(j)_k}{j!}(x - a)^{j-k},$$

where $(j)_k$ is defined to be $j(j-1)(j-2)\cdots(j-k+1)$ ("j to the power k falling"). Evaluating at $x = a$ then gives that the kth derivative at a is $f^{(k)}(a)$.

Taylor polynomial of f at a of order n Suppose f is a function defined at and near a. The *Taylor polynomial of f at a of order n* is

$$P_{n,a,f}(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n$$

We give some examples now, the details of which are left as exercises:

- $P_{n,0,\exp}(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots + \frac{x^n}{n!}$.
- $P_{2n+1,0,\sin}(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots + (-1)^n \frac{x^{2n+1}}{(2n+1)!}$.
- $P_{2n,0,\cos}(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots + (-1)^n \frac{x^{2n}}{(2n)!}$.
- $P_{n,1,\log}(x) = (x-1) - \frac{(x-1)^2}{2} + \frac{(x-1)^3}{3} - \frac{(x-1)^4}{4} + \cdots + (-1)^n \frac{(x-1)^n}{n}$.

The Taylor polynomial is not always easy to calculate. For example, consider $P_{n,0,\tan}$. We have

- $\tan 0 = 0$,
- $\tan' = \sec^2$, so $\tan' 0 = 1$,
- $\tan'' = (\sec^2)' = 2\sec^2\tan$, so $\tan'' 0 = 0$,
- $\tan''' = (2\sec^2\tan)' = 2\sec^4 + 4\sec^2\tan^2$, so $\tan''' 0 = 2$,

and so $P_{3,0,\tan} = x + x^3/3$, but it does not seem very easy to continue.

14.2 Properties of the Taylor polynomial

An important property of the linearization $P_{1,a,f}$ of a function f at a (formerly denoted $L_{f,a}$) is that not only does $P_{1,a,f}(x) - f(x) \to 0$ as $x \to a$, but also

$$\frac{P_{1,a,f}(x) - f(x)}{x-a} = f'(a) - \frac{f(x)-f(a)}{x-a} \to 0 \quad \text{as} \quad x \to a$$

So the error in using $P_{1,a,f}$ to approximate f not only gets smaller as x gets closer to a, but gets smaller relative to $x-a$. But it is not necessarily the case that $(P_{1,a,f}(x)-f(x))/((x-a)^2)$ goes to zero as x approaches a. Consider, for example, the function $f(x) = x^2$ at $a = 0$, for which

$$\frac{P_{1,0,f}(x) - f(x)}{(x-a)^2} = -1 \not\to 0 \quad \text{as} \quad x \to a.$$

What about limit as $x \to a$ of

$$\frac{P_{2,a,f}(x) - f(x)}{(x-a)^2} = \frac{f(a) + f'(a)(x-a) + f''(a)(x-a)^2/2 - f(x)}{(x-a)^2}?$$

By L'Hôpital's rule, this limit is same as

$$\lim_{x\to a}\frac{f'(a)+f''(a)(x-a)-f'(x)}{2(x-a)},$$

if this limit exists; but

$$\frac{f'(a)+f''(a)(x-a)-f'(x)}{2(x-a)}=\frac{1}{2}\left(f''(a)-\frac{f'(x)-f'(a)}{x-a}\right),$$

which approaches 0 as $x\to a$, by the definition of the second derivative at a; so the original limit is 0. Note that $f(x)=x^3$, $a=0$, shows $(P_{2,a,f}(x)-f(x))/((x-a)^3)$ does not necessarily tend to 0.

This example leads to the following definition.

Definition of functions agreeing to order n A function g *agrees* with a function f *to order* n ($n\geq 0$ an integer) *at* a if both g and f are defined near a and if

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

exists and equals 0.

We use the shorthand $g\sim_{n,a}f$ to denote that g agrees with f to order n at a.

Note that if $g\sim_{n,a}f$ then automatically $f\sim_{n,a}g$, so it is legitimate to say "f and g agree to order n at a". Note also that if $f\sim_{n,a}g$ then $f\sim_{m,a}g$ for all $0\leq m<n$, since

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

although it is not necessarily the case that $f\sim_{m,a}g$ for any $m>n$ (as some earlier examples show). Finally, note that if $f\sim_{n,a}g$ and if $g\sim_{n,a}h$ then it follows that $f\sim_{n,a}g$. Indeed:

$$\frac{f(x)-h(x)}{(x-a)^n}=\frac{f(x)-g(x)}{(x-a)^n}+\frac{g(x)-h(x)}{(x-a)^n},$$

the right-hand side above tends at 0 as $x\to a$, so the left-hand side does also.

The example that lead to this definition strongly suggests that the Taylor polynomial $P_{n,a,f}$ agrees with f to order n. That's the content of the next theorem.

Theorem 14.1. *Suppose f is a function such that each of $f,f',f'',\ldots,f^{(n)}$ exist at a. Then $P_{n,a,f}\sim_{n,a}f$ (but is is not necessarily the case that $P_{n,a,f}\sim_{n+1,a}f$).*

Proof: Consider $f(x)=x^{n+1}$ at $a=0$ to see that we may not have agreement to order $n+1$.

For first part we want to show that

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

Set

$$P_n(x) = f(a) + f'(a)(x-a) + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n - f(x)$$

and $Q_n(x) = (x-a)^n$. It is an easy check that of the following limits exist and equal 0:

- $\lim_{x\to a} P_n(x), \lim_{x\to a} P_n'(x), \ldots, \lim_{x\to a} P_n^{(n-2)}(x)$ and
- $\lim_{x\to a} Q_n(x), \lim_{x\to a} Q_n'(x), \ldots, \lim_{x\to a} Q_n^{(n-2)}(x)$.

So, applying L'Hôpital's rule $n-1$ times, we get that the limit in $(\star)$ exists if

$$\lim_{x\to a} \frac{P_n^{(n-1)}(x)}{Q_n^{(n-1)}(x)} = \lim_{x\to a} \frac{f^{(n-1)}(a) + f^{(n)}(a)(x-a) - f^{(n-1)}(x)}{n!(x-a)}$$

exists; but

$$\lim_{x\to a} \frac{f^{(n-1)}(a) + f^{(n)}(a)(x-a) - f^{(n-1)}(x)}{n!(x-a)} = \frac{f^{(n)}(a)}{n!} - \frac{1}{n!} \lim_{x\to a} \frac{f^{(n-1)}(x) - f^{(n-1)}(a)}{(x-a)}$$

which exists and equals 0 by the definition of the nth derivative; so the limit in $(\star)$ exists and equals 0. □

So, the Taylor polynomial of f of degree n at a agrees with f to order n at a. Does this property *characterize* the Taylor polynomial, among polynomials of degree n? Essentially, "yes", as we now see. First we need some notation. Say that Q *is a polynomial of degree at most n in $x-a$* if

$$Q(x) = a_0 + a_1(x-a) + a_2(x-a)^2 + \cdots + a_n(x-a)^n$$

(where a_n is not necessarily non-zero).[229]

Theorem 14.2. *Suppose that f is a function that is n times differentiable at a[230], and that Q is a polynomial of degree at most n in $x-a$ that agrees with f at a to order n. Then $Q = P_{n,a,f}$ (so the degree n Taylor polynomial of f at a is the* unique *polynomial of degree at most n that agrees with f to order n at a).*

[229] For every polynomial of degree m, and for every real a, the polynomial can be expressed as a polynomial of degree m in $x-a$. This is a future-fact — it comes from Linear Algebra. It's easy to see why it is true, though. Here's an example, of expressing an ordinary quadratic polynomial a quadratic polynomial in x) as a quadratic polynomial in $x-1$:

$$x^2 - 4x + 5 = (x-1)^2 - 2x + 4 = (x-1)^2 - 2(x-1) + 2.$$

The trick is to work from the higher powers down.

[230] This hypothesis is necessary. It is *not* true that if f is a function defined at at near a, and if Q is a polynomial of degree at most n in $x-a$ that agrees with f to order n at a, then $Q = P_{n,a,f}$. The issue is that although Q may agree with f to order n at a, it may not be the case that f has the necessary derivatives existing to have a Taylor polynomial. (Spivak gives a specific example in his text.)

This will be a corollary of the following lemma.

Lemma 14.3. *If P and Q are polynomials of degree at most n, and $P \sim_{n,a} Q$, then $P = Q$.*

To see that Theorem 14.2 follows from this, note that

- Q (in the statement of Theorem 14.2) agrees with f to order n at a (by hypothesis of Theorem 14.2),
- $P_{n,a,f}$ agrees with f to order n at a (by Theorem 14.1), so
- Q agrees with $P_{n,a,f}$ to order n at a (as discussed earlier), and so
- $Q = P_{n,a,f}$ (by Lemma 14.3).

Proof (of Lemma 14.3): Set $R = P - Q$, so

$$\frac{R(x)}{(x-a)^n} \to 0 \quad (\star)$$

as $x \to a$. Write $R(x) = r_0 + r_1(x-a) + \cdots + r_n(x-a)^n$.

From $(\star)$ it follows that

$$\frac{R(x)}{(x-a)^i} \to 0 \quad (\star\star)$$

as $x \to a$, for each $i = 0, \ldots, n$.

We have $R(x) \to r_0$ as $x \to a$; but considering $(\star\star)$ at $i = 0$, we get also $R(x) \to 0$ as $x \to a$. So $r_0 = 0$.

From this it follows that $R(x)/(x-a) \to r_1$ as $x \to a$; but considering $(\star\star)$ at $i = 1$, we get $R(x)/(x-a) \to 0$ as $x \to a$. So $r_1 = 0$.

Continuing in this many, we get that $r_i = 0$ for all i, and so $R = 0$ and $P = Q$. □

This theorem suggests an alternate approach to finding Taylor polynomials: if f is n times differentiable at a, and we can someone guess or intuit a polynomial of degree n around a that agrees with f to order n at a, then that polynomial must by the Taylor polynomial of order n at a of f.

Here's an example. Consider $\tanh^{-1} x$ (recall that $\tanh x = (e^x - e^{-x})/(e^x + e^{-x})$), a function with domain $\mathbb{R}$ and range $(-1, 1)$. We know (or can derive) that

$$(\tanh^{-1})'(x) = \frac{1}{1-x^2},$$

so

$$(\tanh^{-1})''(x) = \frac{2x}{(1-x^2)^2} \quad \text{and} \quad (\tanh^{-1})'''(x) = \frac{6x^2+2}{(1-x^2)^3},$$

so $P_{3,0,\tanh^{-1}}(x) = x + x^3/3$. It does not seem like it will be a very pleasant task to continue calculating Taylor polynomials via derivatives!

But we also have

$$\begin{aligned}\tanh^{-1} x &= \int_0^x \frac{dt}{1-t^2} \\ &= \int_0^x \left(1+t^2+t^4+\cdots+t^{2n}+\frac{t^{2n+2}}{1-t^2}\right)\,dt \\ &= x+\frac{x^3}{3}+\frac{x^5}{5}+\cdots+\frac{x^{2n+1}}{2n+1}+\int_0^x \frac{t^{2n+2}}{1-t^2}\,dx. \quad (\star)\end{aligned}$$

If we can show

$$\lim_{x\to 0}\frac{1}{x^{2n+1}}\int_0^x \frac{t^{2n+2}}{1-t^2}dx = 0 \quad (\star\star)$$

then that would exactly say (via $(\star)$) that the polynomial $\sum_{k=0}^n x^{2k+1}/(2k+1)$ agrees with $\tanh^{-1}(x)$ to order $2n+1$ at 0, and so is the Taylor polynomial $P_{2n+1,0,\tanh^{-1}}(x)$.

Using the evenness of the integrand, we have that for $|x|<1/2$

$$\begin{aligned}\left|\int_0^x \frac{t^{2n+2}}{1-t^2}\,dx\right| &= \int_0^{|x|}\frac{t^{2n+2}}{1-t^2}\,dx \\ &\leq \frac{4}{3}\int_0^{|x|} t^{2n+2}dx \\ &= \frac{4|x|^{2n+3}}{3(2n+3)}\end{aligned}$$

and so indeed $(\star\star)$ holds[231] and

$$P_{2n+1,0,\tanh^{-1}}(x) = x+\frac{x^3}{3}+\frac{x^5}{5}+\cdots+\frac{x^{2n+1}}{2n+1}.$$

We can do a little better than this. Let $x\in(-1,1)$ be fixed (note that $(-1,1)$ is the domain of $\tanh^{-1}$). Arguing as above we have

$$\left|\int_0^x \frac{t^{2n+2}}{1-t^2}\,dx\right| \leq \frac{|x|^{2n+3}}{(2n+3)(1-x^2)} \to 0 \quad \text{as } n\to\infty.$$

A corollary of this calculation is that for all $x\in(-1,1)$ we have

$$\left|\tanh^{-1}(x)-P_{2n+1,0,\tanh^{-1}}(x)\right| \leq \frac{4|x|^{2n+3}}{3(2n+3)}.$$

So we can estimate $\tanh^{-1}(x)$ for any particular $x\in(-1,1)$, to any accuracy, by using $P_{2n+1,0,\tanh^{-1}}(x)$ for large enough n. This hints at the main point of what we are about to do: if we can estimate the *difference* between $f(x)$ and $P_{n,a,f}(x)$ (perhaps just for x in some little interval around a), then we can have the potential to use the Taylor polynomial as a reliable way to estimate the function that its the Taylor polynomial of.

[231] Here's another approach: applying L'Hôpital's rule (and the fundamental theorem of calculus) to $(\star\star)$ we find that the limit exists and equals

$$\lim_{x\to 0}\left(\frac{1}{(2n+1)x^{2n}}\right)\left(\frac{x^{2n+2}}{1-x^2}\right) = \lim_{x\to 0}\frac{x^2}{(2n+1)(1-x^2)},$$

as long as these limits exist. But the last limit evidently exists and equals 0.

14.3 Taylor's theorem and remainder terms

Definition of remainder term If f is a function and $P_{n,a,f}$ exists, then the *remainder term* $R_{n,a,f}(x)$ is defined by

$$f(x) = P_{n,a,f}(x) + R_{n,a,f}(x).$$

Our goal for the next while is to find good estimates for $R_{n,a,f}(x)$, that allow us to say that for x sufficiently close to a, $R_{n,a,f}(x) \to 0$ as $n \to \infty$ (so the Taylor polynomial at a is a good approximation for f near a). For example, as we have already discussed, if $x \in [-1/2, 1/2]$ is fixed then

$$R_{2n+1,0,\tanh^{-1}}(x) = \int_0^x \frac{t^{2n+2}}{1-t^2}\ dx \leq \frac{4|x|^{2n+3}}{3(2n+3)} \to 0$$

as $n \to \infty$.[232]

In what follows, we slowly derive Taylor's Theorem with integral remainder term (Theorem 14.4 below). We won't both to mention explicitly the assumptions we are making; those will be stated explicitly in the the theorem, and will easily be seen to be exactly the hypothesis needed to make the argument we are about to describe work.

Let f be a function, with a and x two *fixed* points of the domain of f.[233] We assume $x \neq a$, since it is rather trivial to understand $R_{n,a,f}(a)$. From the fundamental theorem of calculus we have

$$f(x) - f(a) = \int_a^x f'(t)\ dt \quad \text{or} \quad f(x) = f(a) + \int_a^x f'(t)\ dt \text{or} \quad f(x) = P_{0,a,f}(x) + \int_a^x f'(t)\ dt$$

which says that $R_{0,a,f}(x)$ can be expressed as $\int_a^x f'(t)\ dt$.

Now we apply integration by parts to $\int_a^x f'(t)\ dt$, taking

$$\begin{aligned} u = f'(t) \quad &\text{so} \quad du = f''(t)\ dt \\ dv = dt \quad &\text{so} \quad v = t - x. \end{aligned}$$

Notice here that we are *not* taking $v = t$, the obvious choice for an antiderivative of 1. We could, but it would lead us nowhere. Instead we are taking another, non-obvious but equally correct (because x is just some fixed constant), antiderivative; as we will see in a moment, it

[232] Note that we have made a subtle change in viewpoint: we are thinking now of x as being fixed (some number close to a), and thinking about n growing, rather than thinking about n as being fixed with x approaching a.

[233] It is critical that a and x are both consider to be fixed here. Think of a as a point at which we know a lot about f, and of x as some other point, perhaps close to a.

is this non-obvious choice that drives the proof of Taylor's theorem. We get

$$\begin{aligned} f(x) &= f(a) + \int_a^x f'(t)\ dt \\ &= f(a) + [(t-x)f'(t)]_{t=a}^x - \int_a^x (t-x)f''(t)\ dt \\ &= f(a) + (x-a)f'(a) + \int_a^x (x-t)f''(t)\ dt \\ &= P_{1,a,f}(x) + \int_a^x (x-t)f''(t)\ dt. \end{aligned}$$

This says that $R_{1,a,f}(x)$ can be expressed as $\int_a^x (x-t)f''(t)\ dt$.

Now we apply integration by parts to $\int_a^x (x-t)f''(t)\ dt$, taking

$$\begin{aligned} u = f''(t) &\quad \text{so} \quad du = f'''(t)\ dt \\ dv = (x-t)dt &\quad \text{so} \quad v = \frac{-(x-t)^2}{2}. \end{aligned}$$

Notice here that we *are* taking the obvious choice for antiderivative of 1; as we will in all subsequent applications in this proof. We get

$$\begin{aligned} f(x) &= f(a) + (x-a)f'(t) + \int_a^x (x-t)f''(t)\ dt \\ &= f(a) + (x-a)f'(t) + \left[\frac{-(x-t)^2}{2}f''(t)\right]_{t=a}^x + \int_a^x \frac{(x-t)^2}{2}f'''(t)\ dt \\ &= f(a) + (x-a)f'(t) + \frac{(x-a)^2 f''(a)}{2} + \int_a^x \frac{(x-t)^2}{2}f'''(t)\ dt \\ &= P_{2,a,f}(x) + \int_a^x \frac{(x-t)^2}{2}f'''(t)\ dt. \end{aligned}$$

This says that $R_{2,a,f}(x)$ can be expressed as $\int_a^x \frac{(x-t)^2}{2}f'''(t)\ dt$.

We try this one more time. We apply integration by parts to $\int_a^x \frac{(x-t)^2}{2}f'''(t)\ dt$, taking

$$\begin{aligned} u = f'''(t) &\quad \text{so} \quad du = f''''(t)\ dt \\ dv = \frac{(x-t)^2}{2}\ dt &\quad \text{so} \quad v = \frac{-(x-t)^3}{3!}. \end{aligned}$$

We get

$$\begin{aligned} f(x) &= + \int_a^x \frac{(x-t)^2}{2}f'''(t)\ dt \\ &= f(a) + (x-a)f'(t) + \frac{(x-a)^2 f''(a)}{2} + \left[\frac{-(x-t)^3}{3!}f'''(t)\right]_{t=a}^x + \int_a^x \frac{(x-t)^3}{3!}f'''(t)\ dt \\ &= f(a) + (x-a)f'(t) + \frac{(x-a)^2 f''(a)}{2} + \frac{(x-a)^3 f'''(a)}{3!} + \int_a^x \frac{(x-t)^3}{3!}f'''(t)\ dt \\ &= P_{3,a,f}(x) + \int_a^x \frac{(x-t)^3}{3!}f'''(t)\ dt. \end{aligned}$$

This says that $R_{3,a,f}(x)$ can be expressed as $\int_a^x \frac{(x-t)^3}{3!} f''''(t)\ dt$.

An obvious pattern is emerging, and we can verify it by induction. Suppose, for some k, we have shown that

$$f(x) = P_{k,a,f}(x) + \int_a^x \frac{(x-t)^k}{k!} f^{(k+1)}(t)\ dt.$$

We apply integration by parts to $\int_a^x \frac{(x-t)^k}{k!} f^{(k+1)}(t)\ dt$, taking

$$\begin{array}{rcl} u = f^{(k+1)}(t) & \text{so} & du = f^{(k+2)}(t)\ dt \\ dv = \dfrac{(x-t)^k}{k!}\ dt & \text{so} & v = -\dfrac{(x-t)^{k+1}}{(k+1)!}. \end{array}$$

We get

$$\begin{aligned} f(x) &= P_{k,a,f}(x) + \int_a^x \frac{(x-t)^k}{k!} f^{(k+1)}(t)\ dt \\ &= P_{k,a,f}(x) + \left[\frac{-(x-t)^{k+1}}{(k+1)!} f^{(k+1)}(t) \right]_{t=a}^x + \int_a^x \frac{(x-t)^{(k+1)}}{(k+1)!} f^{(k+2)}(t)\ dt \\ &= P_{k,a,f}(x) + \frac{(x-a)^{k+1}}{(k+1)!} f^{(k+1)}(t) + \int_a^x \frac{(x-t)^{(k+1)}}{(k+1)!} f^{(k+2)}(t)\ dt \\ &= P_{k+1,a,f}(x) + \int_a^x \frac{(x-t)^{(k+1)}}{(k+1)!} f^{(k+2)}(t)\ dt. \end{aligned}$$

This says that $R_{k+1,a,f}(x)$ can be expressed as $\int_a^x \frac{(x-t)^{(k+1)}}{(k+1)!} f^{(k+2)}(t)\ dt$.

We have proven the following important theorem:

Theorem 14.4. (Taylor's theorem with integral remainder term) *Suppose $f, f', \ldots, f^{(n+1)}$ are all defined on an interval that includes a and x, and that $f^{(n+1)}$ is integrable on that interval. Then*

$$f(x) = P_{n,a,f}(x) + \frac{1}{n!} \int_a^x (x-t)^n f^{(n+1)}(t)\ dt.$$

That is,

$$R_{n,a,f}(x) = \frac{1}{n!} \int_a^x (x-t)^n f^{(n+1)}(t)\ dt.$$

There is another form of the remainder term that is usually much easier to work with. Suppose that $f^{(n+1)}(t)$ is continuous on the closed interval I that has a and x as endpoints[234]. Then, by the extreme value theorem, there are numbers $m < M$ such that

$$m \leq f^{(n+1)}(t) \leq M$$

[234] We write this, rather than the more natural "on the interval $[a, x]$", to allow for the possibility that $x < a$.

for $t \in I$, and moreover there are numbers $t_1, t_2 \in I$ with $f^{(n+1)}(t_1) = m$ and $f^{(n+1)}(t_2) = M$. If $x > a$ we have that on I

$$m(x-t)^n \leq (x-t)^n f^{(n+1)}(t) \leq M(x-t)^n$$

so, integrating,

$$\frac{m(x-a)^{n+1}}{(n+1)!} \leq R_{n,a,f}(x) \leq \frac{M(x-a)^{n+1}}{(n+1)!}$$

or

$$m \leq \frac{(n+1)! R_{n,a,f}(x)}{(x-a)^{n+1}} \leq M.$$

By the intermediate value theorem, there is some c between t_1 and t_2 (and so between a and x) with

$$f^{(n+1)}(c) = \frac{(n+1)! R_{n,a,f}(x)}{(x-a)^{n+1}}$$

or

$$R_{n,a,f}(x) = \frac{f^{(n+1)}(c)(x-a)^{n+1}}{(n+1)!}.$$

We can use a similar argument to reach the same conclusion, when $x < a$. We summarize in the following theorem, as important as Theorem 14.4.

Theorem 14.5. (Taylor's theorem with Largrange remainder term, weak form[235]) *Suppose $f, f', \ldots, f^{(n+1)}$ are all defined on an interval that includes a and x, and that $f^{(n+1)}$ is continuous on that interval. Then there is some number c (strictly) between a and x such that*

$$f(x) = P_{n,a,f}(x) + \frac{f^{(n+1)}(c)(x-a)^{n+1}}{(n+1)!}.$$

That is,

$$R_{n,a,f}(x) = \frac{f^{(n+1)}(c)(x-a)^{n+1}}{(n+1)!}.$$

14.4 Examples

Example 1, sin at 0 We illustrate the use of Theorem 14.5 with the example of the function $f(x) = \sin x$, at $a = 0$. Fix $x \in \mathbb{R}$. Recall that we have

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + R_{2n+1,0,\sin}(x).$$

The Lagrange form of the remainder term is

$$|R_{2n+1,0,\sin}(x)| = \left| \frac{\sin^{(2n+2)}(c) x^{n+1}}{(n+1)!} \right| \leq \frac{|x|^{2n+2}}{(2n+2)!},$$

[235]Why is this the *weak form*? Because this theorem is also true, without the hypothesis that $f^{(n+1)}$ is continuous. However, since in every example that we will see, we will have continuity of the $(n+1)$st derivative, we will not discuss the stronger form here.

where c is some number between 0 and x. The inequality above follows from the fact that $|\sin^{(2n+2)}(c)| = |\sin(c)| \leq 1$, regardless of the values of n and c.

The continue the analysis, we need the following lemma, that will be extremely useful for many other applications, that says that the factorial function grows faster than any power function.

Lemma 14.6. *For each $x > 0$ and $\varepsilon > 0$, for all sufficiently large n we have*

$$\frac{x^n}{n!} < \varepsilon.$$

Proof: Pick any integer $n_0 > 2x$. We have that for $n > n_0$,

$$\frac{x^n}{n!} \leq \frac{(n_0/2)^n}{n_0^{n-n_0}(n_0-1)!} = \frac{n_0^{n_0}}{(n_0-1)!}\frac{1}{2^n}.$$

Noting that $(n_0^{n_0}/(n_0-1)!)$ is just a constant, we can make $1/(2^n) < \varepsilon/(n_0^{n_0}/(n_0-1)!)$, so $x^n/n! < \varepsilon$, for all sufficiently large n.

Alternately: if n is even, then the largest $n/2$ terms in the product $n!$ are all bigger than $n/2$, so $n! > (n/2)^{n/2}$ while if n is odd, then the largest $(n+1)/2$ terms in $n!$ are all bigger than $n/2$, so $n! > (n/2)^{(n+1)/2}$. Either way, for all n

$$n! > \left(\frac{n}{2}\right)^{n/2} = \left(\frac{\sqrt{n}}{\sqrt{2}}\right)^n,$$

so

$$\frac{x^n}{n!} < \left(\frac{x\sqrt{2}}{\sqrt{n}}\right)^n.$$

For all $n \geq 8x^2$ we therefore have

$$\frac{x^n}{n!} < \left(\frac{1}{2}\right)^n.$$

Since $(1/2)^n$ can be made smaller than ε by choosing n sufficiently large, so too can $x^n/n!$.

□

An immediate corollary is that for each real x,

$$R_{2n+1,0,\sin}(x) \to 0 \quad \text{as } n \to \infty,$$

and so for each real x

$$P_{2n+1,0,\sin}(x) \to \sin x \quad \text{as} \quad n \to \infty.$$

For example, to estimate $\sin 355$ to within $\pm.001$, we simply choose n large enough that $355^{2n+2}/((2n+2)!) < .001$, and then calculate $P_{2n+1,0,\sin}(355)$. A `Mathematica` calculation tells us that $n = 483$ is sufficiently large, and that $P_{2(483)+1,0,\sin}(355) = -0.000233397\ldots$, so we conclude

$$\sin 355 = -0.000233397\ldots \pm 0.001.$$

(In fact $\sin 355 = -0.0000301444\cdots$[236]).

Given that $\sin 355$ is so close to zero, it is rather remarkable that the sequence $(P_{2n+1,0,\sin}(355))_{n\geq 0}$ starts $(355, -7\times 10^6, 4\times 10^{10}, \ldots)$, and that along the way to the term $P_{967,0,\sin}(355)$, the sequence rises up to as large as 1.6×10^{152}!

Example 2, $\cos$ at 0 By an almost identical argument to the one used for sin, we find that for all real x,

$$P_{2n,0,\cos}(x) \to \cos x \quad \text{as} \quad n \to \infty.$$

Example 3, exp We have, for each fixed x,

$$P_{n,0,\exp}(x) = 1 + x + \frac{x^2}{2!} + \ldots + \frac{x^n}{n!}$$

with (Lagrange form of the remainder, c between 0 and x)

$$R_{n,0,\exp}(x) = \frac{\exp(c)}{(n+1)!}x^{n+1}$$

so that, using the fact that exp is an increasing function.

$$|R_{n,0,\exp}(x)| \leq e^{\max\{0,x\}}\frac{|x|^{n+1}}{(n+1)!}$$

From Lemma 14.6 we get that

$$R_{n,0,\exp}(x) =\!\!\to 0 \quad \text{as } n \to \infty$$

and, as with sin, this is valid for *all* real x, so the Taylor polynomial of exp at 0 can be used to estimate $\exp^x$ to arbitrary precision for all x; that is,

$$P_{n,0,\exp}(x) \to \exp x \quad \text{as} \quad n \to \infty.$$

As an illustrative example we estimate e^{-1}. Setting $x = -1$ we have

$$|R_{n,0,\exp}(-1)| \leq \frac{1}{(n+1)!}.$$

[236] Why is this so close to zero? It's because 355 is almost an integer multiple of *pi*; in fact, $355 \approx 113\pi = 354.9999698556467\ldots$. That begs the question, "why is 355/113 such a good approximation to π? That gets into the theory of continued fractions.

Since $1/11! = 0.000000025\ldots$ we get that

$$P_{10,0,\exp}(-1) = \frac{16481}{44800} = 0.367879464\ldots$$

is an approximation of $1/e$ accurate to ± 0.000000025. (In fact $1/e = 0.367879441\ldots$.)

Example 4, $f(x) = e^{-1/x^2}$ at 0 Consider the function f defined by

$$f(x) = \begin{cases} e^{-1/x^2} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0. \end{cases}$$

We studied this function in some detail, as an example of the application of Lemma 12.6. We proved there that f is differentiable arbitrarily many times at 0, and that all derivatives at 0 are 0. It follows that

$$P_{n,0,f}(x) = 0$$

for all n (for all x; $P_{n,0,f}$ is the identically 0 polynomial). It follows that

$$R_{n,0,f}(x) = f(x)$$

for all n and x. So: for $x = 0$ we have $R_{n,0,f}(x) \to 0$ as $n \to \infty$ (trivially), but for $x \neq 0$ we have

$$R_{n,0,f}(x) = e^{-x^2/2} \not\to 0 \quad \text{as } n \to \infty.$$

In this example, the Taylor polynomial is useless as an approximation tool.

Example 5, $\tan^{-1}$ at 0 Consider $\tan^{-1} : \mathbb{R} \to (-\pi/2, \pi/2)$. Proceeding as we did for $\tanh^{-1}$, we have

$$\begin{aligned} \tan^{-1}(x) &= \int_0^x \frac{1}{1+t^2} dt \\ &= x - \frac{x^3}{3} + \cdots + \frac{(-1)^n x^{2n+1}}{2n+1} + \int_0^x \frac{(-1)^{n+1} t^{2n+2}}{1+t^2} dt. \end{aligned}$$

Is $P_{2n+1,0,\tan^{-1}}(x) = \sum_{k=0}^n \frac{(-1)^k x^{2k+1}}{2k+1}$? Yes, because for all x,

$$\begin{aligned} \left| \frac{1}{x^{2n+1}} \int_0^x \frac{(-1)^{n+1} t^{2n+2}}{1+t^2} dt \right| &= \frac{1}{|x|^{2n+1}} \int_0^{|x|} \frac{t^{2n+2}}{1+t^2}\, dt \\ &\leq \frac{1}{|x|^{2n+1}} \int_0^{|x|} t^{2n+2}\, dt \\ &= \frac{x^2}{2n+3} \end{aligned}$$

which goes to 0 as x goes to 0, and so the degree $2n+1$ polynomial we have found agrees with $\tan^{-1}$ to order $2n+1$ at 0.

The calculation we have just done shows that

$$|R_{2n+1,0,\tan^{-1}}(x)| \leq \frac{|x|^{2n+3}}{2n+3}.$$

As long as $x \in [-1,1]$, we therefore have $R_{2n+1,0,\tan^{-1}}(x) \to 0$ as $n \to \infty$, and so in this range $P_{2n+1,0,\tan^{-1}}(x) \to \tan^{-1}(x)$.

What about when $|x| > 1$? We claim that here, $|R_{2n+1,0,\tan^{-1}}(x)| \not\to 0$. We first consider positive x. If $|R_{2n+1,0,\tan^{-1}}(x)| \to 0$ in this case, then, since

$$\begin{aligned}\left|\int_1^x \frac{(-1)^{n+1}t^{2n+2}}{1+t^2}\,dt\right| &= \left|\int_0^x \frac{(-1)^{n+1}t^{2n+2}}{1+t^2}\,dt - \int_0^1 \frac{(-1)^{n+1}t^{2n+2}}{1+t^2}\,dt\right| \\ &\leq \left|\int_0^x \frac{(-1)^{n+1}t^{2n+2}}{1+t^2}\,dt\right| + \left|\int_0^1 \frac{(-1)^{n+1}t^{2n+2}}{1+t^2}\,dt\right|\end{aligned}$$

and each of $\int_0^x \frac{(-1)^{n+1}t^{2n+2}}{1+t^2}\,dt$, $\int_0^1 \frac{(-1)^{n+1}t^{2n+2}}{1+t^2}\,dt \to 0$ as $n \to \infty$, we would have

$$\int_1^x \frac{(-1)^{n+1}t^{2n+2}}{1+t^2}\,dt \to 0.$$

But now, on the interval $[1,x]$ we have

$$\frac{(-1)^{n+1}t^{2n+2}}{1+t^2} \geq \frac{1}{1+x^2}$$

so that

$$\text{either} \quad \int_1^x \frac{(-1)^{n+1}t^{2n+2}}{1+t^2}\,dt \geq \frac{1}{1+x^2} \quad \text{or} \quad \int_1^x \frac{(-1)^{n+1}t^{2n+2}}{1+t^2}\,dt \leq \frac{-1}{1+x^2}$$

(depending on whether n is odd or even), and so it cannot possibly be that the integral tends to 0 as n grows. We conclude that

$$P_{2n+1,0,\tan^{-1}}(x) \to \tan^{-1}(x) \quad \textit{only} \text{ on the interval } [-1,1].$$

The upshot of these examples, is that it seems that for each f and $a \in \text{Domain}(f)$, there is a range of x around a for which, for each fixed x in that range, $R_{n,a,f}(x)$ goes to 0 as n goes to infinity, and so for which $P_{n,a,f}(x)$ approaches $f(x)$ as n gets large.

- In the case of $\sin, \cos, \exp$ at 0, that range is all of $\mathbb{R}$.
- In the case of $\tanh^{-1}$ at 0, that range is $(-1,1)$, which coincides with the domain of $\tanh^{-1}$.
- In the case of $\tan^{-1}$ at 0, that range is goes from -1 to 1 which again includes only a portion of the domain.

- In the case of e^{-1/x^2} at 0, that range includes *only* the point 0.

That $P_{n,0,\exp}(x)$ approaches e^x as n gets large, for all real x, suggests that we can meaningfully write something like

$$\begin{aligned} e^x &= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} \cdots \\ &= \sum_{n=0}^{\infty} \frac{x^n}{n!} \end{aligned}$$

for all $x \in \mathbb{R}$, and that $P_{n,0,\tan^{-1}}(x)$ approaches $\tan^{-1} x$ as n gets large, for all $x \in (-1, 1)$ suggests that the equation

$$\tan^{-1}(x) = x - \frac{x^3}{3} + \frac{x^5}{5} - \cdots$$

is meaningful for all $x \in [-1, 1]$.

The goal of rest of these notes is to study infinite sequences and series, to make what we have just discussed precise.