A simple motivation for Taylor polynomial approximation

One motivation behind the concept of Taylor polynomials is to find better and better approximations to the value \( f(x) \) of a ‘nice’ function \( f \), when the number \( x \) is close to a number \( a \) where the values of the function and its derivatives are easy to compute. So suppose \( f \) and \( a \) are given, and we can easily compute \( f(a) \), \( f'(a) \), \( f''(a) \), etc. And suppose that \( x \) is near \( a \), but \( f(x) \) is not so easy to compute. As a first approximation, we could say

\[ f(x) \approx f(a) \]

since \( x \approx a \). Sometimes this is ok — great precision is not always needed.

But, how can one improve? The strategy we follow is simple: Use the fundamental theorem of calculus to write an exact version of the approximate equation. This will include an error term expressed as an integral of \( f' \). Then approximate the error term by recursively applying the same strategy, to \( f'(t_1) \approx f'(a) \), \( f''(t_2) \approx f''(a) \), etc.

1. At first, let’s just see where this strategy takes us. In the first few steps we find

\[ f(x) = f(a) + \int_a^x f'(t_1) \, dt_1 \]

\[ = f(a) + \int_a^x f'(a) \, dt_1 + \int_a^x \int_a^{t_1} f''(t_2) \, dt_2 \, dt_1 \]

\[ = f(a) + \int_a^x f'(a) \, dt_1 + \int_a^x \int_a^{t_1} f''(a) \, dt_2 \, dt_1 + \int_a^x \int_a^{t_1} \int_a^{t_2} f'''(t_3) \, dt_3 \, dt_2 \, dt_1. \]

We don’t know much about the last term at this point, but for the other terms, notice

\[ \int_a^x f'(a) \, dt_1 = f'(a) \int_a^x \, dt_1 = f'(a)(x-a), \]

\[ \int_a^x \int_a^{t_1} f''(a) \, dt_2 \, dt_1 = f''(a) \int_a^x (t_1-a) \, dt_1 = f''(a) \frac{(x-a)^2}{2}, \]

\[ \int_a^x \int_a^{t_1} \int_a^{t_2} f'''(a) \, dt_3 \, dt_2 \, dt_1 = f'''(a) \int_a^x \frac{(t_1-a)^2}{2} \, dt_1 = f'''(a) \frac{(x-a)^3}{3 \cdot 2}, \]

and in general

\[ \int_a^x \int_a^{t_1} \ldots \int_a^{t_n-1} f^{(n)}(a) \, dt_n \ldots dt_2 \, dt_1 = f^{(n)}(a) \frac{(x-a)^n}{n!}. \]

One conjectures, then, and proves by induction, that

\[ f(x) = P_n(x) + R_n(x) \]

where \( P_n \) is the \( n \)-th Taylor polynomial

\[
P_n(x) = f(a) + f'(a)(x-a) + f''(a)\frac{(x-a)^2}{2} + \ldots + f^{(n)}(a)\frac{(x-a)^n}{n!}
\]

and the \( n \)-th Taylor remainder \( R_n(x) \) is represented by nested integrals as

\[ R_n(x) = \int_a^x \int_a^{t_1} \ldots \int_a^{t_n} f^{(n+1)}(t_{n+1}) \, dt_{n+1} \ldots dt_2 \, dt_1. \]

This works provided, e.g., that \( f^{(j)} \) is continuously differentiable for \( j = 1, \ldots, n \).
2. The form of the remainder \( R_n(x) \) above looks complicated, but actually it is not difficult to prove excellent bounds for it, if bounds on \( f^{(n+1)} \) are known. (Remark: This is the main way Taylor’s theorem gets used.) Consider the case \( x > a \) first. Let \( m_{n+1} \) be the minimum value of \( f^{(n+1)} \) on \([a, x]\), and \( M_{n+1} \) the maximum value. Then since

\[
m_{n+1} \leq f^{(n+1)}(t_{n+1}) \leq M_{n+1} \quad \text{for all } t_{n+1} \text{ in } [a, x],
\]

after \( n + 1 \) repeated integrations as above one finds, for \( x > a \),

\[
m_{n+1} \frac{(x-a)^{n+1}}{(n+1)!} \leq R_n(x) \leq M_{n+1} \frac{(x-a)^{n+1}}{(n+1)!}.
\]

If \( x < a \) and \( n \) is odd, the same bounds hold. If \( x < a \) and \( n \) is even, \((x-a)^{n+1} < 0\) and the bounds hold with the inequalities reversed.

These bounds show that there are several ways the remainder can be guaranteed tiny; then the Taylor polynomials lead to excellent approximations of \( f(x) \). First, as \( x \) approaches \( a \), the factor \((x-a)^{n+1} \) in the error bounds goes to zero faster than the last nonvanishing term in \( P_n(x) \), which is proportional to \((x-a)^n \) at worst. So increasing \( n \) improves the order of the approximation near \( a \). Second, the impressive factor \((n+1)! \) in the denominator grows extremely fast as \( n \) becomes large. It will start to make the ratio \((x-a)^{n+1}/(n+1)! \) smaller as soon as \( n + 1 > |x-a| \), in fact, no matter how big \(|x-a| \) is. (How good the approximation is overall depends on how fast the bounds \( m_{n+1} \) and \( M_{n+1} \) grow with \( n \).) And third, it can happen that the bounds \( m_{n+1} \) and \( M_{n+1} \) become small as \( n \) gets large (which means high derivatives of \( f \) become smaller as the order of differentiation increases).

3. A more compact way to represent the remainder (called the Lagrange form) emerges by applying the intermediate and extreme value theorems, using simple comparisons as follows. Notice that the function

\[
t \mapsto f^{(n+1)}(t) \frac{(x-a)^{n+1}}{(n+1)!}
\]

attains the extreme values

\[
m_{n+1} \frac{(x-a)^{n+1}}{(n+1)!} \quad \text{and} \quad M_{n+1} \frac{(x-a)^{n+1}}{(n+1)!}
\]

at some points between \( x \) and \( a \) (presuming continuity of \( f^{(n+1)} \)). By the intermediate value theorem, there must be some point \( t \) between these two points—so between \( x \) and \( a \)—such that

\[
R_n(x) = f^{(n+1)}(t) \frac{(x-a)^{n+1}}{(n+1)!}.
\]

This is the Lagrange form of the remainder.

4. Another fairly convenient form for the remainder \( R_n(x) \) is called the integral form:

\[
R_n(x) = \int_a^x f^{(n+1)}(t) \frac{(x-t)^n}{n!} \, dt
\]

(It is significant that this formula also works for functions \( f \) whose values are \textit{vectors}.) One proves this formula directly by induction using integration by parts, keeping careful track of sign:

\[
\int_a^x f^{(n+1)}(t) \frac{(x-t)^n}{n!} \, dt = f^{(n+1)}(t) \frac{(x-t)^{n+1}}{(n+1)!} \bigg|_{t=a}^{t=x} + \int_a^x f^{(n+2)}(t) \frac{(x-t)^{n+1}}{(n+1)!} \, dt.
\]