Taylor polynomials

© 2012 Yonatan Katznelson

1. *Introduction*

The most elementary functions are *polynomials* because they involve only the most basic arithmetic operations of addition and multiplication. Polynomials are also easy to differentiate, and their long term behavior is also very easy to understand.

Mathematical modeling of economic phenomena, however, often leads to functions which are not polynomials, like exponential functions and logarithm functions, and combinations of functions that involve division, extracting roots, etc.

Under certain conditions, it is possible to find polynomials that provide good *approximations* to more general functions. In the following sections I'll outline one of the most basic and important ways of doing this for functions that are differentiable. I'll focus on the first and second degree approximations (linear and quadratic), but for completeness' sake, I'll briefly describe the general case as well.

2. *Linear approximation.*

For a function $f(x)$ that is differentiable at a point $x = x_0$, we deduced the following approximation formula from the definition of the derivative,

$$f(x) - f(x_0) \approx f'(x_0)(x - x_0). \tag{2.1}$$

I mentioned that this approximation is accurate when the difference $|x - x_0|$ is *sufficiently small.* Adding $f(x_0)$ to both sides of (2.1), gives a formula for approximating $f(x)$ in the neighborhood of x_0 by a linear function,[†]

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0). \tag{2.2}$$

The linear function, $T(x) = f(x_0) + f'(x_0)(x - x_0)$, should look familiar to you, because its graph is precisely the tangent line to the graph $y = f(x)$ at the point $(x_0, f(x_0))$, and the approximation (2.2) has a simple geometric interpretation. Namely, the tangent line $y = T(x)$ is ***close*** to the graph $y = f(x)$, when x is sufficiently close to x_0. By close, I mean that the ***vertical distance***, $|T(x) - f(x)|$, is small.

Example 1. The graph of the function $f(x) = \sqrt{-x^2 + 4x + 25} - 2$, and the graph of the linear approximation to this graph at the point $(4, 3)$, $T(x) = 3 - 0.4(x - 4)$, are both displayed in Figure 1.

A quick glance at the figure shows two things. First, the tangent line is very close to the graph when x is close to 4. Second, as x moves away from 4 (in either direction), the

[†]In mathematical terms, a '*neighborhood*' of x_0 is an interval around x_0 of the form $(x_0 - \delta, x_0 + \delta)$, where δ is a small, positive constant.

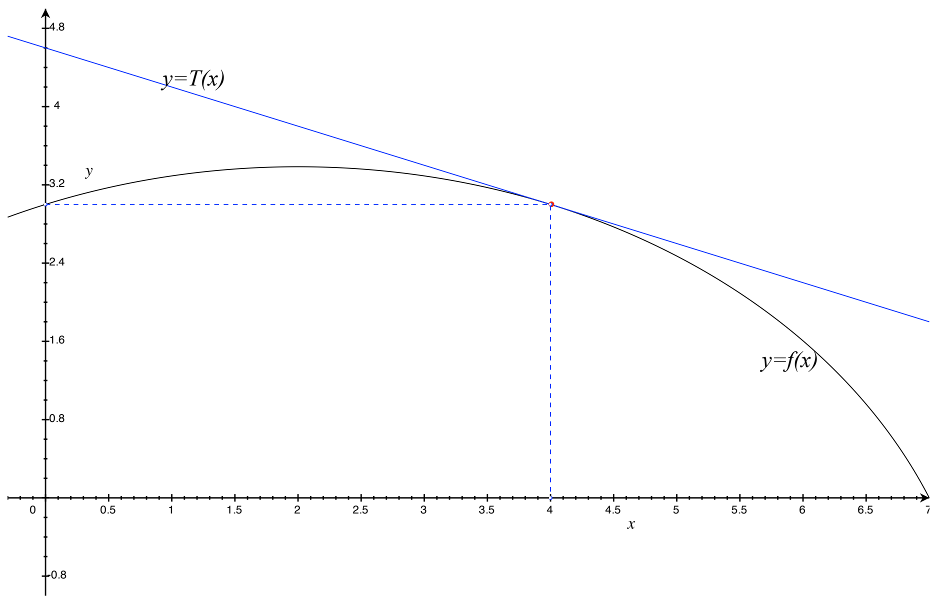

Figure 1: Linear approximation to $f(x) = \sqrt{-x^2 + 4x + 25} - 2$, centered at $(4, 3)$.

vertical distance between the tangent line and the graph grows larger. The explanation for this first phenomenon has already been given,[‡] but we can also explain it as follows.

The two functions, $T(x)$ and $f(x)$, have the same value and the same slope at $x_0 = 4$, i.e., $T(4) = f(4)$ and $T'(4) = f'(4)$. This means that their graphs emanate from the same point $(4, 3)$ and they both move in the same direction (same slope). So for a while, the two graphs stay close to each other.

The explanation for the second phenomenon is also relatively simple. The slope of the tangent line is constant (equal to $f'(4)$), but the slope of the graph $y = f(x)$ is *not constant*. As x moves away from 4, the derivative of $f(x)$ changes. And, as the slope changes, the graph $y = f(x)$ 'bends away' from the tangent line $y = T(x)$.

3. *The second order Taylor approximation.*

The rate of change of the derivative $f'(x)$ is given by the second derivative, $f''(x)$. This fact suggests the following idea: *To improve upon the linear approximation to the function $f(x)$ near the point x_0, look for a function that has the same value at x_0 as $f(x)$, the same derivative at x_0 as $f(x)$, **and the same second derivative** at x_0 as $f(x)$. Furthermore, the function we're looking for should be as simple as possible.*

In other words, we're looking for a (differentiable) function $T_2(x)$ satisfying

$$T_2(x_0) = f(x_0), \quad T_2'(x_0) = f'(x_0) \text{ and } T_2''(x_0) = f''(x_0). \tag{3.1}$$

And $T_2(x)$ should be 'simple'. The simplest functions are linear, but if $T_2(x) = ax + b$, then $T_2'(x) = a$ and $T_2''(x) = 0$. If $f''(x_0) = 0$, then the linear approximation fits the bill, but if $f''(x_0) \neq 0$, then the function that we're looking for *cannot* be a linear function.

[‡]Earlier in the course.

The next simplest functions are ***quadratic functions***. If $T_2(x) = ax^2 + bx + c$, then $T_2'(x) = 2ax + b$ and $T_2''(x) = 2a$. The conditions in (3.1) give us a system of three equations for a, b and c. Namely

$$\begin{aligned} ax_0^2 + bx_0 + c &= f(x_0) \\ 2ax_0 + b &= f'(x_0) \\ 2a &= f''(x_0). \end{aligned} \tag{3.2}$$

This system of linear equations is easy to solve,[§] but by being a little clever, we can make our work even easier. The idea is to express the (as of yet unknown) function $T_2(x)$ in terms of the *difference* $(x - x_0)$. We are still looking for a quadratic function, but we write it as

$$T_2(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)^2.$$

The motivation for this idea comes from the linear approximation, which is naturally expressed in terms of $(x - x_0)$.

The first and second derivatives of $T_2(x)$, written in this way, are $T_2'(x) = a_1 + 2a_2(x - x_0)$ and $T_2''(x) = 2a_2$. When we impose the conditions (3.1) on this version of $T_2(x)$ and its derivatives, the simple system (3.2) is replaced by the *very* simple system

$$\begin{aligned} a_0 + a_1(x_0 - x_0) + a_2(x_0 - x_0)^2 &= f(x_0) \\ a_1 + 2a_0(x_0 - x_0) &= f'(x_0) \\ 2a_2 &= f''(x_0), \end{aligned} \tag{3.3}$$

which is very simple because $x_0 - x_0 = 0$, so the system (3.3) reduces to

$$a_0 = f(x_0), \quad a_1 = f'(x_0) \text{ and } a_2 = \frac{f''(x_0)}{2}.$$

Definition 1.

The ***second degree Taylor polynomial for f(x) centered at x_0*** *is the quadratic function*

$$T_2(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2. \tag{3.4}$$

This function satisfies the conditions

$$T_2(x_0) = f(x_0), \quad T_2'(x_0) = f'(x_0) \text{ and } T_2''(x_0) = f''(x_0).$$

The approximation

$$f(x) \approx T_2(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 \tag{3.5}$$

is called the ***second order Taylor approximation***. The second order approximation is usually better than the linear (first order) approximation, as long as the difference $|x - x_0|$ is small enough.

[§] Try it, starting with solving the third equation for a, then working backwards to solve for b and then c.

Example 2. Returning to the function $f(x) = \sqrt{-x^2 + 4x + 25} - 2$ from Example 1, we find that it's second degree Taylor polynomial, centered at $(4, 3)$, is given by

$$T_2(x) = 3 - 0.4(x - 4) - 0.116(x - 4)^2,$$

as you can (and should!) verify by direct computation.

The graphs of $f(x)$, its linear approximation and its quadratic approximation are all displayed in Figure 2, with the quadratic approximation in red.

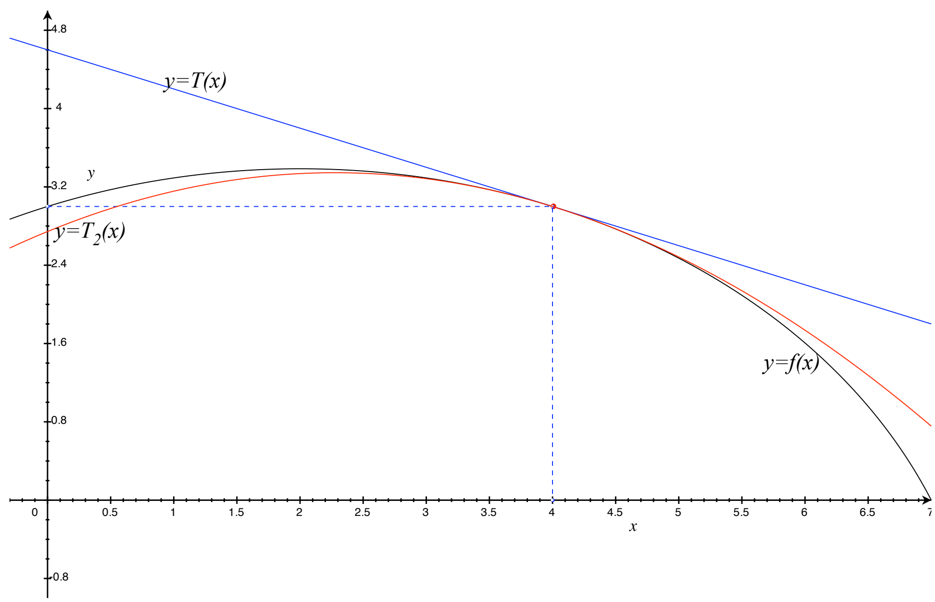

Figure 2: Linear and quadratic approximations to $f(x) = \sqrt{-x^2 + 4x + 25} - 2$.

As you can see, the quadratic approximation is better than the linear approximation in two ways. First, the quadratic approximation is closer to the graph of the original function than the tangent line. Second, it stays closer to the original graph for longer. This second phenomenon can be explained by the fact that the quadratic approximation is *bending* in the same way as the original graph around the point $(4, 3)$. More properly said, $T_2(x)$ and $f(x)$ *have the same concavity* at the point $(4, 3)$.

To get a numerical view of how the second order approximation improves on the linear approximation, consider the next example.

Example 3. Let $f(x) = \sqrt{x} = x^{1/2}$, and let $x_0 = 16$. The first and second derivatives of this function are $f'(x) = \frac{1}{2}x^{-1/2}$ and $f''(x) = -\frac{1}{4}x^{-3/2}$. Thus,

$$f(16) = 4, \quad f'(16) = \frac{1}{8} \text{ and } f''(16) = -\frac{1}{256}.$$

The first degree (linear) Taylor polynomial for $f(x) = x^{1/2}$, centered at $x = 16$ is therefore given by

$$T(x) = 4 + \frac{1}{8}(x - 16)$$

and the second degree Taylor polynomial (centered at the same point) is given by

$$T_2(x) = 4 + \frac{1}{8}(x-16) - \frac{1}{512}(x-16)^2.$$

I'll test the accuracy of the linear and quadratic approximations for $x = 16.81$, for which we have $\sqrt{16.81} = 4.1$. The linear approximation gives

$$T(16.81) = 4 + \frac{1}{8}(16.81 - 16) = 4.10125,$$

and the quadratic approximation gives

$$T_2(16.81) = 4 + \frac{1}{8}(16.81-16) - \frac{1}{512}(16.81-16)^2 = 4.0999685546875.$$

The linear approximation is off by $|4.10125 - 4.1| = 0.00125$, while the quadratic approximation is off by $|4.0999685546875 - 4.1| = 0.0000314453125$. The quadratic approximation, in this case, is more than 30 times closer to the truth than the linear approximation.

4. *The Taylor polynomial of a function.*

In general, there is no reason to stop at two derivatives. Assume that the function $f(x)$ has derivatives of order up to and including n, defined at the point x_0. The system of equations (3.3), that we solved to find the coefficients of the quadratic Taylor polynomial, generalizes easily to find a polynomial $T_n(x)$ satisfying the $n+1$ conditions

$$\left.\begin{array}{rcl} T_n(x_0) &=& f(x_0), \\ T_n'(x_0) &=& f'(x_0), \\ T_n''(x_0) &=& f''(x_0), \\ T_n'''(x_0) &=& f'''(x_0), \\ &\vdots& \\ T_n^{(n)}(x_0) &=& f^{(n)}(x_0). \end{array}\right\} \tag{4.1}$$

Generalizing what we did in the quadratic case, we write

$$T_n(x) = a_0 + a_1(x-x_0) + a_2(x-x_0)^2 + \cdots + a_n(x-x_0)^n,$$

and use the conditions in (4.1) to obtain a very simple system of equations for the coefficients $a_0, a_1, \ldots, a_n$. Indeed, the first condition gives

$$f(x_0) = T_n(x_0) = a_0 + a_1(x_0-x_0) + a_2(x_0-x_0)^2 + \cdots + a_n(x_0-x_0)^n = a_0,$$

which implies that $a_0 = f(x_0)$. Differentiating once gives

$$T'(x) = a_1 + 2a_2(x-x_0) + 3a_3(x-x_0)^2 + \cdots + na_n(x-x_0)^{n-1},$$

so $T'(x_0) = a_1$ and the second condition, $T'(x_0) = f'(x_0)$, implies that $a_1 = f'(x_0)$. Differentiating again gives

$$T''(x) = 2a_2 + 6a_3(x-x_0) + \cdots + n(n-1)a_n(x-x_0)^{n-2},$$

so $T''(x_0) = 2a_2$ and the condition $T''(x_0) = f''(x_0)$ implies that $a_2 = f''(x_0)/2$. Continuing, we have

$$T'''(x) = 6a_3 + 24a_4(x - x_0) + \cdots + n(n-1)(n-2)a_n(x - x_0)^{n-3},$$

so $T_n'''(x_0) = 6a_3$, and the condition $T'''(x_0) = f'''(x_0)$ implies that $a_3 = f'''(x_0)/6$.

In general, Differentiating k times ($k \le n$) and evaluating at $x = x_0$, gives $T^{(k)}(x_0) = (1 \cdot 2 \cdot 3 \cdots k) \cdot a_k$, and the condition $T_n^{(k)}(x_0) = f^{(k)}(x_0)$ implies that

$$a_k = \frac{f^{(k)}(x_0)}{k!},$$

where $k!$, pronounced '***k factorial***' is shorthand for the product $1 \cdot 2 \cdot 3 \cdots k$.¶

Definition 2.
The polynomial

$$T_n(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 + \cdots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n \qquad (4.2)$$

*is called the n^{th} degree **Taylor polynomial** of $f(x)$, centered at $x = x_0$. This is the unique polynomial of degree n that satisfies the $n + 1$ conditions in* (4.1).

Example 4. Let's find the 4^{th} degree Taylor polynomial of $f(x) = \ln x$, centered at $x = 1$. First, we compute the derivatives up to and including order 4:

$$f'(x) = x^{-1}, \quad f''(x) = -x^{-2}, \quad f'''(x) = 2x^{-3} \text{ and } f^{(4)}(x) = -6x^{-4}.$$

Evaluating $\ln x$ and its derivatives at $x = 1$, and using the definition of $T_4(x)$, above, we find that

$$\begin{aligned} T_4(x) &= \ln 1 + 1^{-1} \cdot (x-1) + \frac{-1^{-2}}{2} \cdot (x-1)^2 + \frac{2 \cdot 1^{-3}}{6}(x-1)^3 + \frac{-6 \cdot 1^{-4}}{24}(x-1)^4 \\ &= (x-1) - \frac{1}{2}(x-1)^2 + \frac{1}{3}(x-1)^3 - \frac{1}{4}(x-1)^4. \end{aligned}$$

The graphs of $y = \ln x$ and $y = T_4(x)$ (red) appear in Figure 3. These graphs highlight two important features of the Taylor polynomial and how well it may or may not approximate the function from which it was derived.

First of all, looking at the whole picture shows that the Taylor polynomial $T_4(x)$ behaves completely differently than $\ln x$. For example, as we can easily see in the figure, if $x > 3$, then $T_4(x) < -1$ (and $T_4(x)$ is decreasing), while $\ln x > 1$ (and $\ln x$ is increasing).

On the other hand, if we only look in the vicinity of the point $x = 1$, the two functions are almost identical. In Figure 3, the two graphs are indistinguishable when $|x - 1| < 1/2$.‖

¶This function grows very rapidly, e.g., $1! = 1$, $2! = 2$, $3! = 6$, $4! = 24$, $5! = 120$, $10! = 3628800$ and $15! = 1307674368000$. In fact, when n is very large, $n! \approx (n/e)^n \cdot \sqrt{2\pi n}$. This approximation is called *Stirling's formula*.

‖The graphing utility that I use rounds plot positions to 4 decimal places, so points on the two graphs that are less than 0.0005 apart may appear to coincide.

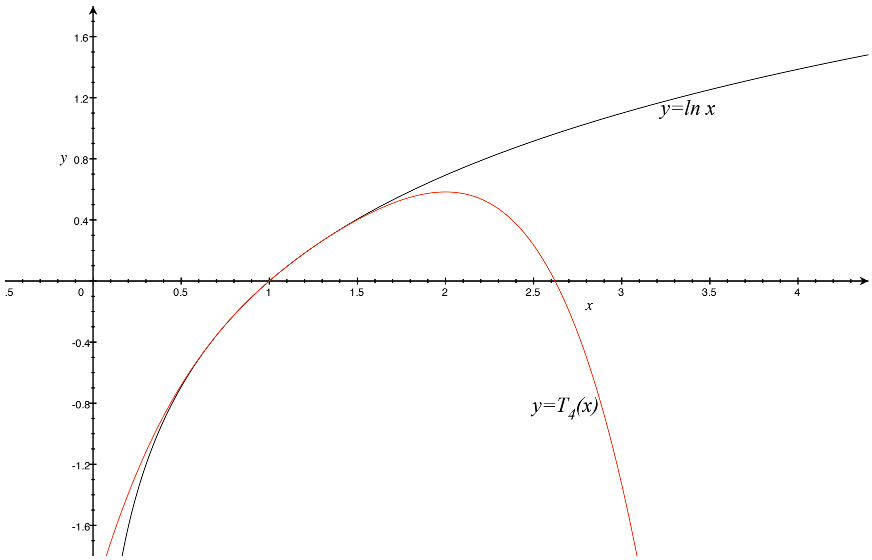

Figure 3: The graphs of $\ln x$ and $T_4(x)$.

If you prefer numerical evidence, you can evaluate $\ln x$ and $T_4(x)$ on your favorite calculator, at points close to 1, and see how far apart the values are. According to my TI-30XA

- $\ln 1.5 - T_4(1.5) \approx 0.0044234$,
- $\ln 1.25 - T_4(1.25) \approx 0.0001618$ and
- $\ln 1.1 - T_4(1.1) \approx 0.0000018$.

The errors of approximation have all been rounded to 7 decimal places, and as you can see, the closer the argument is to 1, the better the approximation.

Exercises.

1. Compute the 2^{nd} degree Taylor polynomial for the function $R(x) = \dfrac{x^3 + 2x + 1}{x + 3}$ centered at the point $(1, 1)$. Use the quadratic Taylor approximation to estimate $R(1.2)$. What is the error of the approximation?

2. Find the 3^{rd} degree Taylor polynomial for the function $f(x) = \sqrt{x}$, centered at $x_0 = 100$. Use this to compute an approximate value for $\sqrt{110}$.

3. Find the 4^{th} degree Taylor polynomial for the function $g(x) = e^x$, centered at $x_0 = 0$. Use this to compute an approximate value for $\sqrt[4]{e} = e^{1/4}$.

4. a. Find the 7^{th} degree Taylor polynomial for $f(x) = \ln x$, centered at $x_0 = 1$.

b. Use your answer to a. to find approximate values for $\ln(2/3)$ and $\ln(3/4)$.

c. Use your answers to b. to find an approximate value for $\ln(3)$. ***Do not use your answer to part a.*** (Hint: $3 = (4/3) \cdot (3/2) \cdot (3/2)$, $3/2 = 1/(2/3)$ and $4/3 = 1/(3/4)$. Use properties of the logarithm function.)