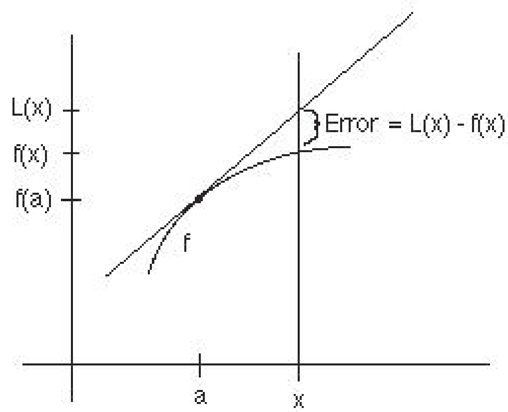
Linear Approximations

One of the most important ideas associated with the tangent line at a point on the graph of a function is that it provides a local linearization of the function. That is, no matter how complicated the graph of a differentiable function, no matter the difficulty of computing function values, near a point where the derivative exists, we can approximate the function by the tangent line. For this reason we have the following definition.

Definition: Let f be a differentiable function. Then the linearization (or linear approximation) of f about $x = a$ is the function $L(x)$ defined by

$$L(x) = f(a) + f'(a)(x - a)$$

Note that L is the linear function whose graph is the tangent line at $(a, f(a))$. In particular, instead of values on the graph of f near $(a, f(a))$, we use values on the tangent line, which may be relatively more straightforward to compute. Of course, there usually will be some error because the tangent line is different from the graph of the function. However, $\lim_{x\to a}(L(x) - f(x)) = f(a) - f(a) = 0$ implies that that $L(x) \approx f(x)$ for values of x near a.



Example 1: Find the linearization of $f(x) = \sqrt{x}$ about $x = 9$. Use it to approximate $\sqrt{8}$. In this example, $f'(x) = \frac{1}{2\sqrt{x}}$. Thus,

$$\begin{aligned} L(x) &= \sqrt{9} + \frac{1}{2\sqrt{9}}(x - 9) \\ &= 3 + \frac{1}{6}(x - 9) \\ L(8) &= 3 + \frac{1}{6}(8 - 9) \\ &= \frac{17}{6} \end{aligned}$$

Example 2: Use an appropriate linerization to find an approximate value of $\cos(36\,\text{deg})$. We change to radians: $36\,\text{deg} = (36/180)\pi = \pi/5$ radians. Thus, to approximate $\cos(\pi/5)$ we let $f(x) = \cos x$ and

$a = \pi/6$. Hence,

$$\begin{aligned} L(x) &= \cos\frac{\pi}{6} + \left(-\sin\frac{\pi}{6}\right)\left(x - \frac{\pi}{6}\right) \\ &= \frac{\sqrt{3}}{2} - \frac{1}{2}\left(x - \frac{\pi}{6}\right) \\ L(\frac{\pi}{5}) &= \frac{\sqrt{3}}{2} - \frac{1}{2}\left(\frac{\pi}{5} - \frac{\pi}{6}\right) \end{aligned}$$

The concept of linearization is especially important because it is this notion of the derivative as the *best linear approximation* that generalizes to functions of more than one variable. The linearization is *best* in the sense that not only does $\lim_{x\to a}(L(x) - f(x)) = 0$, but also

$$\lim_{x\to a} \frac{L(x) - f(x)}{x - a} = 0$$

That is, $L(x) \to f(x)$ faster than $x \to a$. We will close by showing that the above limit is indeed zero:

$$\begin{aligned} \lim_{x\to a} \frac{L(x) - f(x)}{x - a} &= \lim_{x\to a} \frac{f(a) + f'(a)(x - a) - f(x)}{x - a} \\ &= \lim_{x\to a} \left(f'(a) - \frac{f(x) - f(a)}{x - a}\right) \\ &= f'(a) - f'(a) \\ &= 0 \end{aligned}$$

Modeling note: If f has a derivative at $x = a$, we have defined the linearization of f about $x = a$ to be the function L defined by $L(x) = f(a) + f'(a)(x - a)$. The function L is actually a linear polymial in x which we can rewrite as $P_1(x) = f(a) + f'(a)(x - a)$ to emphasize this fact. This polynomial agrees with f and f' at $x = a$: $P_1(a) = f(a)$ and $P_1'(a) = f'(a)$. It is also the *best* linear approximation of f near a.

Suppose now that the derivatives $f^{(j)}(x)$ exist on the interval I for $j = 1, 2, 3 \cdots n$, where $f^{(1)}(x) = f'(x)$, $f^{(2)}(x) = f''(x)$, $f^{(3)}(x) = f'''(x)$, etc. If a is an interior point of I, define P_n on I as

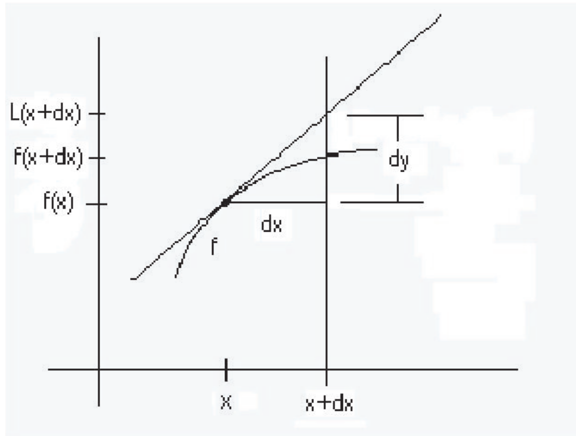
$$P_n(x) = f(a) + \frac{f'(a)}{1!}(x - a) + \frac{f''(a)}{2!}(x - a)^2 + \frac{f'''(a)}{3!}(x - a)^3 + \cdots + \frac{f^{(n)}(a)}{n!}(x - a)^n$$

Then it is clear that $P_n(a) = f(a)$, $P_n^{(j)}(a) = f^{(j)}(a)$ for $j = 1, 2, 3 \cdots n$, and we state without proof that P_n is the best approximation of f around $x = a$ of any polynomial of degree n.

This is a version of Taylor's Theorem, and $P_n(x)$ is called the Taylor polynomial of f of degree n.When we first studied the derived table for $y = e^x$, we pointed out that if the underlying function were not known, linear, quadratic, and cubic polynomials might suffice to model the data dependinging on what else was known about the real-world problem from which the data came. The polynomials we gave were simply the corresponding Taylor polynomials for e^x of degrees 1, 2, and 3; indeed, they gave better and better fits of the data as the degree increased.

Example 3: Find the taylor polynomials of degree 1, 2, and 3 for $f(x) = e^x$ around $x = 0$. Use them to approximate $e^{0.02}$. The derivatives of f are $f^{(n)}(x) = e^x$ for $n = 1, 2, 3$. Hence, $f^{(n)}(0) = 0$ for all n. Thus, $P_1(x) = 1 + x$, $P_2(x) = 1 + x + \frac{1}{2}x^2$, and $P_3(x) = 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3$. Hence, substituting, we get $P_1(0.02) = 1 + .02 = 1.02$, $P_2(0.02) = 1 + .02 + .5(.02)^2 = 1.02020$, and $P_3(0.02) = 1 + .02 + .5(.02)^2 + \frac{1}{6}(.02)^3 = 1.020201333$. These should be compared with the Maple value $e^{0.02} \approx 1.020201340$.

Differentials: Consider now a slight modification of the sketch with which we started the section.



The new sketch shows the tangent line to the graph of a differentiable function $y = f(x)$ at the point (x, y). Now let dx be an increment in x, and let dy be the resulting vertical rise of the tangent line as shown. Then the ratio of dy to dx equals the slope of the tangent line, $f'(x)$.

$$\frac{dy}{dx} = f'(x)$$

Note that in this equation, dy/dx is the ratio of two numbers, not a notation for the derivative.

Until now, dx and dy could not be separated from the symbol $\frac{dy}{dx}$. When we separate them in the way we just did, we call dx the *differential of* x and dy the *differential of* y. Formally, given a differentiable function $y = f(x)$, we treat dx as an independent variable, and define the dependent variable dy according to the formula $dy = f'(x)dx$. That is, dy is obtained by the formal multiplication of $f'(x)$ and dx.

For example, if $y = x^2$, then $dy = 2x\,dx$. Or if $y = \ln x$, then $dy = \frac{1}{x}dx$.

Leibniz had the idea that the notation of calculus should facilitate its use. The new meanings that we have given to the symbols dy and dx do just that, but we will not experience their full power until we put them to use when we learn to integrate in Chapter 3.

Relating differentials to the linear approximation $L(x)$ that we discussed above, note that if $\Delta x = dx$ is a change in x at the point (x, y), then the corresponding change Δy in y is $\Delta y = f(x + dx) - f(x)$, while $dy = L(x + dx) - L(x)$. Since $L(x) = f(x)$, we have that $dy = L(x + dx) - f(x)$. Thus, in general Δy and dy are different and the absolute error in using the tangent line approximation near (x, y) is $|\Delta y - dy| = |f(x + dx) - L(x + dx)|$. Hence, we can think of the differential dy as an approximation to Δy.

Applet: Best Linear Approximation **Try it!**

Exercises: Problems **Check what you have learned!**

Videos: Tutorial Solutions **See problems worked out!**