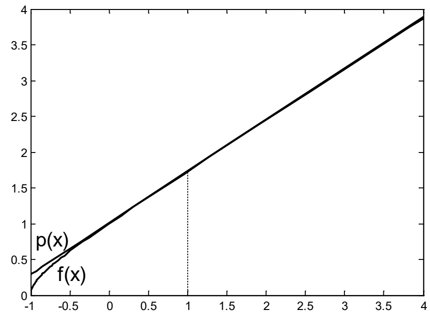
12. Linear Approximations

Linear functions are very easy to work with. For instance, it is a lot easier to solve a set of linear equations, than to solve a set of nonlinear equations. Besides the applications we have seen so far, another application of derivatives is in obtaining 'linear approximations' to nonlinear functions.

Suppose we are working with the function

$$f(x) = (1 + 3x/2 + x^2/2)^{1/2}$$

and that we are interested in particular with the function for values of x near $x = 1$.



The fact that $f(x)$ is differentiable at $x = 1$ means that for values of x near $x = 1$, the tangent line to $f(x)$ at $x = 1$ is a good approximation to $f(x)$. We call the tangent to $f(x)$ at $x = 1$ the "linear approximation" to $f(x)$ at the point $x = 1$, and if we are only interested in the function around the neighborhood of $x = 1$, this approximation may be sufficient for our purposes. That is, it might be sufficient for us to work with the tangent, rather than the original function. From the figure above, it appears that the tangent to the function at $x = 1$ is a good approximation to the actual function for a fairly large neighborhood around that point.

What is the formula for the tangent? Let $p(x) = a + bx$ represent the tangent to a function $f(x)$ at the point $x = x_0$. This tangent can be defined as the line that satisfies

$$\text{(i)} \quad p(x_0) = f(x_0) \quad \text{and} \quad \text{(ii)} \quad p'(x_0) = f'(x_0).$$

That is, the value of the tangent at x_0 is the same as the value of the function at x_0, and the slope of the tangent at x_0 is the same as the slope of $f(x)$ at x_0. We have

$$p(x_0) = a + bx_0 = f(x_0) \qquad \text{and} \qquad p'(x_0) = b = f'(x_0)$$

Solving for a and b, and substituting into $p(x) = a + bx$ gives

$$p(x) = f(x_0) + f'(x_0)(x - x_0).$$

In other words, $p(x)$ is a line with slope $f'(x_0)$ and passes through the point $(x_0, f(x_0))$. Therefore

$$\frac{p(x) - f(x_0)}{x - x_0} = f'(x_0)$$

which gives the same formula.

Example 12.1 Find the linear approximation of

$$f(x) = (1+x)^m \text{ at } x = 0 .$$

We have $f(0) = (1+0)^m = 1$. Also, $f'(x) = m(1+x)^{m-1}$, so $f'(0) = m$. Therefore the linear approximation at $x = 0$ is

$$p(x) = f(0) + f'(0)(x-0) = 1 + mx .$$

Consider the case $m = 1/3$. For values of x near 0, the tangent at $x = 0$ gives good approximations to the actual function around that point:

(i) $\sqrt[3]{1.1} = 1.032280115456367....$ Using the linear approximation $p(x)$ with $m = 1/3$ and $x = 1/10$ gives $1 + (1/10)/3 = 1 + 1/30 = 1.0\dot{3}$.

(ii) $\sqrt[3]{1.01} = 1.003322283542089....$ Using the linear approximation $p(x)$ with $m = 1/3$ and $x = 1/100$ gives $1 + (1/100)/3 = 1 + 1/300 = 1.00\dot{3}$.

Example 12.2 Find an approximate value for $\sqrt{37}$.

Because $\sqrt{37} = \sqrt{36+1} = 6\sqrt{1+1/36}$, we can use the result in Example 12.1.1. We have $6\sqrt{1+x} \approx 6p(x) = 6(1+x/2)$. At $x = 1/36$, we have

$$\sqrt{37} \approx 6(1+(1/36)/2) = 6 + 1/12 = 6.08\dot{3}.$$

The actual value is $\sqrt{37} = 6.082762530298219...$

Exercise Find the linear approximation of the function

$$f(x) = (1 + 3x/2 + x^2/2)^{1/2} \text{ at the point } x = 1 .$$

The following is an application from statistics.

Example 12.3 Recall from elementary statistics that if X is a random variable with some mean and variance, then the variance of $aX + b$ is simply $\text{var}[aX+b] = a^2 \text{var}[X]$. The variance of more complicated function of X is more difficult to derive, for example, what is the variance of

$$f(X) = (1 + 3X/2 + X^2/2)^{1/2} ?$$

Taking the linear approximation of $f(X)$ around a suitable point simplifies the problem substantially.

Differentials The idea of linear approximations can be expressed in terms of changes. Let $y = f(x)$ be the function to be approximated. Consider a small change in x from $x = x_0$ to $x = x_0 + dx$. The actual change in f is $f(x_0 + dx) - f(x_0)$. Suppose we wish to use the linear approximation to the function at x_0 to approximate this change. This approximate change would be $p(x_0 + dx) - p(x_0)$. We have

$$\begin{aligned} p(x_0 + dx) - p(x_0) &= [f(x_0) + f'(x_0)(x_0 + dx - x_0)] - [f(x_0) + f'(x_0)(x_0 - x_0)] \\ &= [f(x_0) + f'(x_0)\,dx] - [f(x_0)] \\ &= f'(x_0)dx \end{aligned}$$

For small values of dx we can expect $p(x_0 + dx) - p(x_0)$ to be a good approximation to $f(x_0 + dx) - f(x_0)$. We denote the linear approximation to the actual change as "df", and write

$$df = p(x_0 + dx) - p(x_0) = f'(x_0)\, dx$$

We can apply this argument to any x, and so we can write in general

$$df = f'(x)\, dx$$

The symbols 'df' and 'dx' are called the differentials of f and x respectively. If we write the function as $y = f(x)$, we can write $dy = f'(x)dx$.

<u>Example 12.4</u> For $f(x) = (1+x)^{1/3}$, we have $f'(x) = (1/3)(1+x)^{-2/3}$. The differential is thus

$$df = f'(x)dx = \frac{1}{3}(1+x)^{-\frac{2}{3}}\, dx$$

E.g., for a small change in x from $x = 0$ to $x = 0.01$, the linear approximation to the change in the value of $f(x)$ would be $f'(0)\, 0.01 = (1/3)(0.01) = 0.0\dot{3}$.

<u>Example 12.5</u> The volume of a sphere of radius r is $V(r) = \frac{4}{3}\pi r^3$, and $V'(r) = 4\pi r^2$. The differential is therefore

$$dV = 4\pi r^2\, dr$$

If radius increases from 2 to 2.03, the linear approximation to the change in volume is

$$dV = 4\pi \underbrace{(2)}_{r}{}^2\ \underbrace{0.03}_{dr} = 1.508.$$

The actual change in volume is $\frac{4}{3}\pi(2.03^3) - \frac{4}{3}\pi 2^3 = 1.531$.

Differentials vs Derivatives The derivative of a function, and the differential of a function are intimately related, but *they are not the same thing*. The derivative of a function tells us the slope of the function at any given point. It is a function of x. The differential is a formula for computing linear approximations to actual changes in f for a given change in x by dx. It is a function of x and dx.

The formula for the differential $dy = f'(x)dx$ can be viewed as the reason behind the dy/dx notation for derivatives (which is sometimes called Liebniz's notation for the derivative). Note, however, that we use the derivative to define the differential, not the other way around.

Although the two are different objects, the equations $dy / dx = 2x$ and $dy = 2x\,dx$ nonetheless do contain the same information about the function. Sometimes differentiation is done in the language of differentials. For instance, you might see a statement like

"differentiating $y = x^2 \exp(x)$ gives $dy = (2+x)\,x \exp(x)\,dx$."

We will revisit all this later, when we discuss derivatives and differentials for functions of many variables.

Quadratic Approximations Suppose we try to approximate f near $x = a$ using a quadratic function instead of a linear function. Suppose we choose that quadratic function $q(x)$ such that

(i) $q(a) = f(a)$, (ii) $q'(a) = f'(a)$, and (iii) $q''(a) = f''(a)$

We would usually write quadratic equations in the form

$$q(x) = c_1 x^2 + c_2 x + c_3 .$$

However, for this purpose, it is convenient to use an alternative form:

$$q(x) = A(x-a)^2 + B(x-a) + C .$$

The two are equivalent, since

$$\begin{aligned} q(x) &= A(x-a)^2 + B(x-a) + C \\ &= A(x^2 - 2ax + a^2) + B(x-a) + C \\ &= Ax^2 + (B - 2Aa)x + (Aa^2 - Ba + C) \end{aligned}$$

The benefit of the alternate form for our application will be obvious from the following computations. We want to find A, B, and C. We have

$$q'(x) = 2A(x-a) + B \quad \text{and} \quad q''(x) = 2A ,$$

which gives (i) $q(a) = C$,

(ii) $q'(a) = 2A(a-a) + B = B$, and

(iii) $q''(a) = 2A$.

Therefore $C = f(a)$, $B = f'(a)$, and $2A = f''(a)$, i.e., $A = f''(a)/2$. The desired quadratic approximation is therefore

$$f(x) \approx q(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2.$$

We call this the *second order Taylor polynomial approximation* of $f(x)$ at $x = a$.

Example 12.6 Find the quadratic approximation of $f(x) = (1+x)^m$ at $x = 0$.

We have $f(0) = (1+0)^m = 1$;

$f'(x) = m(1+x)^{m-1}$, so $f'(0) = m$; and

$f''(x) = m(m-1)(1+x)^{m-2}$, so $f''(0) = m(m-1)$

Therefore the quadratic approximation at $x = 0$ is

$$p(x) = f(0) + f'(0)(x-0) + \frac{f''(0)}{2}(x-0)^2 = 1 + mx + \frac{m(m-1)}{2}x^2$$

For values of x near 0, this gives good approximations:

(i) $\sqrt[3]{1.1} = 1.032280115456367...$. Using the quadratic approximation $p(x)$ with $m = 1/3$ and $x = 1/10$ gives $1 + (1/3)(1/10) + (1/3)(-2/3)(1/2)(1/100) = 1.03\dot{2}$.

(ii) $\sqrt{37} = 6.082762530298219...$ Because $\sqrt{37} = \sqrt{36+1} = 6\sqrt{1+1/36}$, we can use the quadratic approximation $p(x)$ for $\sqrt{1+1/36}$ with $m = 1/2$ and $x = 1/36$, this gives

$$\begin{aligned}\sqrt{37} &\approx 6[1 + (1/36)(1/2) + (1/2)(-1/2)(1/2)(1/36^2)] \\ &= 6.082754629629629...\end{aligned}.$$

Example 12.7 Let $y = f(x)$ be some function. Consider a small change in x from $x = x_0$ to $x = x_0 + \Delta x$. The actual change in f is $f(x_0 + \Delta x) - f(x_0)$. Suppose that instead of computing the actual change in f, we use the quadratic approximation $q(x_0 + dx) - q(x_0)$ as an approximation to the actual change. We have

$$\begin{aligned}q(x_0+\Delta x) - q(x_0) &= [f(x_0) + f'(x_0)(x_0 + \Delta x - x_0) + \frac{f''(x_0)}{2}(x_0 + \Delta x - x_0)^2] - f(x_0) \\ &= f'(x_0)\Delta x + \frac{f''(x_0)}{2}(\Delta x)^2\end{aligned}$$

This is the "quadratic version" of the differential. If we denote this $q(x_0 + \Delta x) - q(x_0)$ by Δf or Δy, we have

Function	$y = f(x)$
Linear approximation	$df = f'(x)dx$
Quadratic approximation	$\Delta y = f'(x)\Delta x + \frac{f''(x)}{2}(\Delta x)^2$

Taylor Series It is straightforward to extend all this further. For example, we can use a cubic polynomial approximation to the function f near $x = a$ by choosing a function $p(x)$ such that

(i) $p(x) = f(a)$ when $x = a$,

(ii) $p'(x) = f'(a)$ when $x = a$,

(iii) $p''(x) = f''(a)$ when $x = a$,

(iv) $p'''(x) = f'''(a)$ when $x = a$.

Following an argument similar to previous sections, we get the equation $p(x)$ to be

$$f(x) \approx p(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2 + \frac{f^{(3)}(a)}{3!}(x-a)^3$$

We call this the *third order Taylor polynomial approximation* of $f(x)$ at $x = a$.

If we repeat this process *n* times we get

$$f(x) \approx f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2 + \frac{f^{(3)}(a)}{3!}(x-a)^3 + \ldots + \frac{f^{(n)}(a)}{n!}(x-a)^n$$

This is called the *nth-order Taylor polynomial approximation* of $f(x)$ at *a*.

It can be shown (but we won't do it!) that the difference between *f*(*x*) and the *n*-th order Taylor approximation is

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

where *c* is a number between *x* and *a*. In other words, we can write

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

for some number c between x and a, as long as the first $n+1$ derivatives exist.

An aside: the special case of the above with $n=0$ is called the Mean-Value Theorem:

$$f(x) = f(a)+f'(c)(x-a) \text{ for some } c \text{ between } x \text{ and } a.$$

Rewriting slightly, we have

$$f'(c) = \frac{f(x)-f(a)}{x-a} \text{ for some } c \text{ between } x \text{ and } a.$$

This says that there is some point c between x and a such that the slope of the function at c is equal to the slope of the line joining the two points $(x, f(x))$ and $(a, f(a))$.

We can use the expression

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

gives us some idea about the size of the approximation error:

Example 12.8 For the function $f(x)=(1+x)^m$, the linear approximation at $x=0$ was found to be $p(x)=1+x/3$. The remainder term is

$$R_2(x)=\frac{f''(c)}{2}(x-0)^2$$

for some c between x and 0. Since $f''(x)=\frac{1}{3}\left(-\frac{2}{3}\right)(1+x)^{-5/3}$, this works out to

$$R_2(x)=-\frac{1}{9}(1+c)^{-5/3}x^2.$$

For x close to zero, c will also be very small, so $(1+c)^{-5/3}$ should be close to one. So, for instance, for $x=0.1$ the error is approximately $-1/900$.

Example 12.9 Let $f(x)=e^x$. Then $f^{(n)}(x)=e^x$ for all *n*. The *n*-th ordered linear approximation for $f(x)$ around the point $x=0$ is

$$\begin{aligned} e^x &\approx f(0)+f'(0)x+\frac{f''(0)}{2}x^2+\frac{f^{(3)}(0)}{3!}x^3+\ldots+\frac{f^{(n)}(0)}{n!}x^n \\ &= 1+x+\frac{1}{2}x^2+\frac{1}{3!}x^3+\ldots+\frac{1}{n!}x^n \end{aligned}.$$

One application of this result is an approximation for the number e. Evaluating the equation at $x=1$ gives

$$e \approx 1+1+\frac{1}{2}+\frac{1}{3!}+\ldots+\frac{1}{n!}.$$

How accurate is the approximation? Or, to turn the question around a little, suppose we wish approximate the number e accurate to 5 decimal places. Then, what order approximation should we use? The error for the n-th order approximation is

$$R_{n+1}(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x-0)^{n+1} = \frac{e^c}{(n+1)!}x^{n+1}$$

for some c between 0 and x. Because we are evaluating the approximation at $x=1$ specifically, this is

$$R_{n+1}(1) = \frac{e^c}{(n+1)!}.$$

But e^x is an increasing function, and $0<c<1$, so the largest value for the error is

$$\frac{e}{(n+1)!} \quad .$$

This may not appear terribly useful, because it uses e, which is what we're trying to approximate in the first place, but we know $e<3$, so we know the error is less that $3/(n+1)!$. For instance, to make

$$\frac{3}{(n+1)!} < 0.000005,$$

we require, $(n+1)! < 3/0.000005 = 600{,}000$, which gives $n>9$. We should take at least 10 terms. For any fixed x the final term goes to zero when we take the limit $n\to\infty$. Thus,

$$e^x = 1+x+\frac{1}{2}x^2+\frac{1}{3!}x^3+\ldots+\frac{1}{n!}x^n+\ldots$$

When $x = 1$, we get the infinite expansion

$$e = 1+1+\frac{1}{2!}+\frac{1}{3!}+\frac{1}{4!}+\ldots$$

Taylor's Series / Maclaurin's Series If we continue the process of taking higher and higher approximations, we 'end up' with the infinite series

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

This expansion is called the Taylor series. When a is set at zero, it is called the Maclaurin series. These are infinite series representations of functions.

Exercises

1. Let $y = f(x) = \sqrt{x}$.

    (a) Find the differential of $f(x)$.

    (b) Compute $\Delta y = f(1.1) - f(1)$. Compute $dy = f'(1)\,dx$ for $dx = 0.1$. What is the error when using dy to approximate Δy.

    (c) Find the linear approximation to $f(x)$ about $x = 1$ (i.e., find the equation of the tangent line to the function $f(x)$ at $x = 1$.) Draw, in a single diagram, both the function and the tangent line at $x = 1$. Mark out in your diagram Δx, Δy, and dy from part (b) (i.e., I am asking you to draw a diagram similar to Figure 2 of pg 218 of your textbook.)

    (d) Compute $\Delta y = f(9.1) - f(9)$. Compute $dy = f'(9)\,dx$ for $dx = 0.1$. What is the error when using dy to approximate Δy.

    (e) Find the linear approximation to $f(x)$ about $x = 9$ (i.e., find the equation of the tangent line to the function $f(x)$ at $x = 9$.) Draw in a single diagram both the function and the tangent line at $x = 9$. Mark out in your diagram Δx, Δy, and dy from part (d)

    (f) Find the quadratic approximation to $f(x)$ at $x = 1$ and at $x = 9$.

2. Find the first, second, third and fourth order approximations to the function

$$f(x) = x^3 + 2x^2 + 2x + 1$$

   at $x = x_0$.

3. Show that the linear approximation to $f(x) = 1/(1+2x)^5$ about $x = 0$ is $1 - 10x$.

4. Write

$$e^x = 1 + x + \frac{1}{2}x^2 + \frac{1}{3!}x^3 + \ldots + \frac{1}{n!}x^n + \ldots$$

using summation notation (remember that $0! = 1$ by definition). Show that $e^{-\lambda}\sum_{i=1}^{\infty}\frac{\lambda^i}{i!} = 1$.