Difference Equations
to
Differential Equations

Section 3.1

Best Affine Approximations

We are now in a position to discuss the two central problems of calculus as mentioned in Section 1.1. In this chapter we will take up the problem of finding tangent lines; in Chapter 4 we will consider the problem of finding areas. We choose this order only because the work we do in solving the tangent line problem in this chapter will be of use, through the Fundamental Theorem of Calculus, in solving area problems in the next.

We begin with some preliminary notation and terminology. If f is a function with domain contained in the set A and range contained in the set B, then we may denote this fact by writing $f : A \to B$. For example, if $g(t) = \sqrt{1-t^2}$ and $\mathbb{R}$ denotes the set of real numbers, then the statements $g : \mathbb{R} \to \mathbb{R}$, $g : [-1,1] \to \mathbb{R}$, and $g : [-1,1] \to [0,1]$ are all correct. We will work exclusively with functions of the form $f : \mathbb{R} \to \mathbb{R}$ until Chapter 7, where we will introduce functions of the form $f : \mathbb{R} \to \mathbb{C}$ and $f : \mathbb{C} \to \mathbb{C}$, where $\mathbb{C}$ denotes the set of complex numbers.

We call a function $f : \mathbb{R} \to \mathbb{R}$ *linear* if there is a constant m such that $f(x) = mx$ for all values of x. Graphically, linear functions are functions whose graphs are straight lines passing through the origin. We call a function $f : \mathbb{R} \to \mathbb{R}$ *affine* if there are constants m and b such that $f(x) = mx + b$ for all values of x. Graphically, affine functions are functions whose graphs are straight lines, not necessarily passing through the origin. Put another way, an affine function is a first degree polynomial. Thus $f(x) = 3x$ is both linear and affine, whereas $g(t) = 4t - 6$ is affine but not linear.

The problem of finding the tangent line for the graph of a given function f at a point (x_0, y_0) is really the problem of finding the affine function T which best approximates f for points close to x_0. In this section we will discuss how to solve this problem. In the remaining sections of this chapter we will consider techniques for finding best affine approximations and discuss some applications. In Chapter 5 we will see how to improve upon affine approximations by using higher degree polynomials.

The following example should help to make these ideas more concrete.

Example Consider the problem of approximating the function $f(x) = \sqrt{x}$ for values of x close to 1. For a first approximation, we might say that

$$x \approx 1$$

for x close to 1. In other words, if we let

$$T(x) = 1$$

for all x, then we are saying that the affine function T is a good approximation for f when x is close to 1. Two facts characterize this statement. First, T and f agree at 1; that is,

$$T(1) = 1 = f(1). \tag{3.1.1}$$

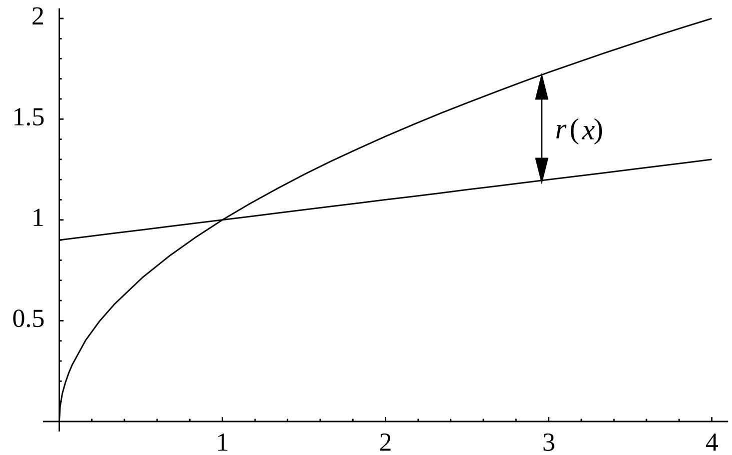

Figure 3.1.1 Graph of $f(x) = \sqrt{x}$ and an approximating affine function

Second, the error committed by approximating f by T goes to 0 as x approaches 1. That is, if we let

$$r(x) = f(x) - T(x),$$

then $r(x)$ is the error made in approximating f by T at the point x, and

$$\lim_{x\to 1} r(x) = \lim_{x\to 1}(f(x) - T(x)) = \lim_{x\to 1}(\sqrt{x} - 1) = 1 - 1 = 0. \tag{3.1.2}$$

Hence we have found an affine function which approximates our function f about $x = 1$ according to some reasonable criterion.

However, it is easy to see that any affine function T whose graph passes through $(1, 1)$ will satisfy (3.1.1) and (3.1.2). First note that if the graph of T is a straight line passing through $(1, 1)$ with slope m, then, using the point-slope form for the equation of a line,

$$T(x) = m(x - 1) + 1.$$

It then follows that

$$T(1) = 1 = f(1)$$

and, if we again let $r(x) = f(x) - T(x)$,

$$\lim_{x\to 1} r(x) = \lim_{x\to 1}(f(x) - T(x)) = \lim_{x\to 1}(\sqrt{x} - (m(x - 1) + 1)) = 1 - 1 = 0.$$

See Figure 3.1.1 for the geometrical interpretation. So now we must ask if there is a value of m which makes T, in some sense, better than any other affine function for approximating f for x near 1. In answering this question, it is convenient to let $h = x - 1$ and to define

$$R(h) = r(1 + h) = f(1 + h) - T(1 + h),$$

the amount of error committed when f is approximated by T at a point a distance h from 1. Since h approaches 0 as x approaches 1, (3.1.1) and (3.1.2) become, in terms of R,

$$R(0) = 0 \tag{3.1.3}$$

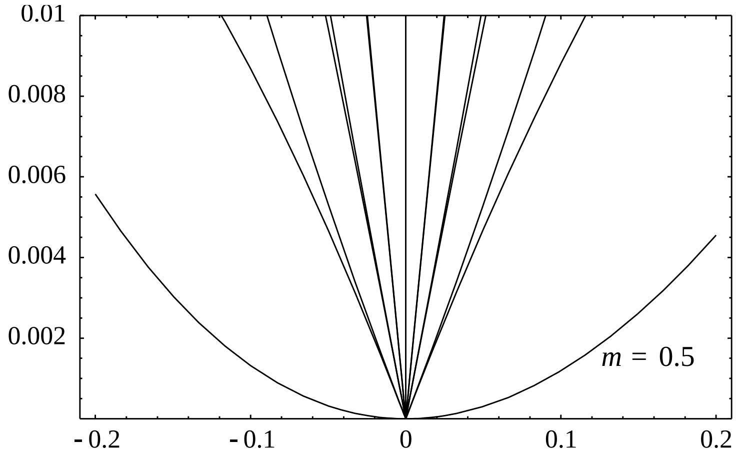

Figure 3.1.2 Graphs of $|R(h)|$ for different values of m

and

$$\lim_{h \to 0} R(h) = 0. \tag{3.1.4}$$

In this case, we have

$$R(h) = \sqrt{1+h} - (m((1+h)-1)+1) = \sqrt{1+h} - (mh+1).$$

Figure 3.1.2 shows the graphs of $|R(h)|$ on the interval $[-0.2, 0.2]$ for $m = 0.1$, 0.3, 0.4, 0.5, 0.6, 0.7, and 0.9. Note that although all these functions approach 0 as h approaches 0, one of the graphs clearly stands out from the others. Namely, when $m = 0.5$, the absolute value of the approximation error appears to approach 0 at a significantly faster rate than does the error for other values of m. To see why this is so, consider that

$$\begin{aligned}
R(h) &= \sqrt{1+h} - (mh+1) \\
&= (\sqrt{1+h} - (mh+1)) \frac{\sqrt{1+h} + (mh+1)}{\sqrt{1+h} + (mh+1)} \\
&= \frac{1+h-(mh+1)^2}{\sqrt{1+h}+mh+1} \\
&= \frac{1+h-(m^2h^2+2mh+1)}{\sqrt{1+h}+mh+1} \\
&= \frac{h(1-2m)-m^2h^2}{\sqrt{1+h}+mh+1}.
\end{aligned}$$

Note that when $m = 0.5$, the numerator reduces to $-m^2h^2$, whereas for other values of m there is also the term $h(1-2m)$. This explains why in Figure 3.1.2 the graph for $m = 0.5$ looks parabolic while the other graphs appear more as straight lines. Moreover, since, for small values of h, h^2 is significantly smaller than h (for example, $(0.001)^2 = 0.000001$ is much smaller than 0.001), we see why the approximation errors when $m = 0.5$ are so much smaller than they are for other values of m.

Intuitively, we should think that for small values of h, $R(h)$ behaves like a multiple of h when $m \neq 0.5$ and like a multiple of h^2 when $m = 0.5$. To see this algebraically, it is useful to consider the quotient

$$\frac{R(h)}{h} = \frac{1 - 2m - m^2 h}{\sqrt{1+h} + mh + 1}.$$

Notice that

$$\lim_{h \to 0} \frac{R(h)}{h} = \frac{1 - 2m}{2},$$

which is 0 only when $m = 0.5$. We interpret this as an indication that $R(h)$ behaves like a multiple of h for small values of h when $m \neq 0.5$, but approaches 0 more rapidly than h when $m = 0.5$.

In our example, we saw that

$$\lim_{h \to 0} \frac{R(h)}{h} = 0$$

when $m = 0.5$, but

$$\lim_{h \to 0} \frac{R(h)}{h} \neq 0$$

for all other values of m. We distinguish the two cases by saying that in the first case $R(h)$ is $o(h)$, whereas in the second case $R(h)$ is only $O(h)$.

Definition A function f is said to be $o(h)$ if

$$\lim_{h \to 0} \frac{f(h)}{h} = 0. \tag{3.1.5}$$

Definition A function f is said to be $O(h)$ if there exist constants M and $\epsilon > 0$ such that

$$\left| \frac{f(h)}{h} \right| \leq M \tag{3.1.6}$$

whenever $-\epsilon < h < \epsilon$.

Note that if

$$\lim_{h \to 0} \frac{f(h)}{h} = L,$$

then we may find an $\epsilon > 0$ such that

$$\left| \frac{f(h)}{h} - L \right| \leq 1$$

whenever $|h| < \epsilon$. Hence

$$L - 1 \leq \frac{f(h)}{h} \leq L + 1$$

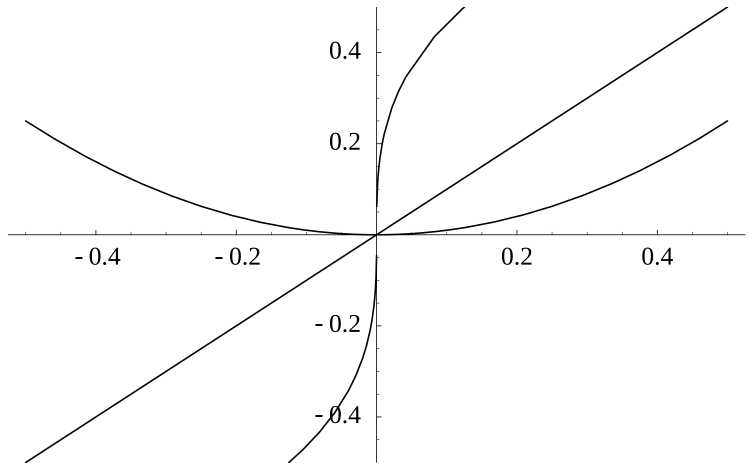

Figure 3.1.3 Rates of convergence to 0 of $f(x) = x^2$, $g(x) = x$, and $k(x) = x^{\frac{1}{3}}$

whenever $|h| < \epsilon$. If we let M be the larger of $|L-1|$ and $|L+1|$, then this shows that

$$\left|\frac{f(h)}{h}\right| \le M$$

whenever $-\epsilon < h < \epsilon$. Hence we have the following proposition.

Proposition If $\lim_{h\to 0} \frac{f(h)}{h}$ exists, then f is $O(h)$.

Note that a function which is $o(h)$ is also $O(h)$. Intuitively, we think of a function which is $o(h)$ as approaching 0 faster than h as h goes to 0, and a function which is $O(h)$ as approaching 0 at a rate which is at least as fast as that of h.

Example Let $f(x) = x^2$, $g(x) = x$, and $k(x) = x^{\frac{1}{3}}$. Then

$$\lim_{h\to 0} f(h) = \lim_{h\to 0} h^2 = 0,$$

$$\lim_{h\to 0} g(h) = \lim_{h\to 0} h = 0,$$

and

$$\lim_{h\to 0} k(h) = \lim_{h\to 0} h^{\frac{1}{3}} = 0.$$

However,

$$\lim_{h\to 0} \frac{f(h)}{h} = \lim_{h\to 0} \frac{h^2}{h} = \lim_{h\to 0} h = 0,$$

so f is $o(h)$;

$$\lim_{h\to 0} \frac{g(h)}{h} = \lim_{h\to 0} \frac{h}{h} = \lim_{h\to 0} 1 = 1,$$

so g is $O(h)$, but not $o(h)$; and

$$\lim_{h\to 0} \frac{k(h)}{h} = \lim_{h\to 0} \frac{h^{\frac{1}{3}}}{h} = \lim_{h\to 0} \frac{1}{h^{\frac{2}{3}}} = \infty,$$

so k is neither $o(h)$ nor $O(h)$. Note in Figure 3.1.3 the difference in the way in which these functions approach 0.

Example Let $f(x) = x - x^2$. Then

$$\lim_{h \to 0} \frac{f(h)}{h} = \lim_{h \to 0} \frac{h - h^2}{h} = \lim_{h \to 0} (1 - h) = 1,$$

so f is $O(h)$, but not $o(h)$.

Example Let $g(x) = 2\sqrt{1+x} - x - 2$. Then

$$\begin{aligned}
\lim_{h \to 0} \frac{g(h)}{h} &= \lim_{h \to 0} \frac{2\sqrt{1+h} - h - 2}{h} \\
&= \lim_{h \to 0} \left(\frac{2\sqrt{1+h} - (h+2)}{h} \right) \left(\frac{2\sqrt{1+h} + (h+2)}{2\sqrt{1+h} + (h+2)} \right) \\
&= \lim_{h \to 0} \frac{4(1+h) - (h+2)^2}{h(2\sqrt{1+h} + (h+2))} \\
&= \lim_{h \to 0} \frac{4 + 4h - (h^2 + 4h + 4)}{h(2\sqrt{1+h} + h + 2)} \\
&= \lim_{h \to 0} \frac{-h^2}{h(2\sqrt{1+h} + h + 2)} \\
&= \lim_{h \to 0} \frac{-h}{2\sqrt{1+h} + h + 2} \\
&= \frac{0}{4} = 0.
\end{aligned}$$

Thus g is $o(h)$.

Example Returning to the problem of approximating $f(x) = \sqrt{x}$ for x close to 1, let

$$T(x) = m(x - 1) + 1$$

and

$$R(h) = f(1 + h) - T(1 + h).$$

We saw above that

$$\lim_{h \to 0} \frac{R(h)}{h} = \frac{1 - 2m}{2}.$$

Thus $R(h)$ is $o(h)$ if and only if $m = 0.5$. In other words, the error in approximating $f(x) = \sqrt{x}$ by the affine function

$$T(x) = \frac{1}{2}(x - 1) + 1 \tag{3.1.7}$$

goes to 0 faster as x approaches 1 than the error for any other affine function approximation. Because of this, we will call (3.1.7) the *best affine approximation* to f at 1. Moreover, we

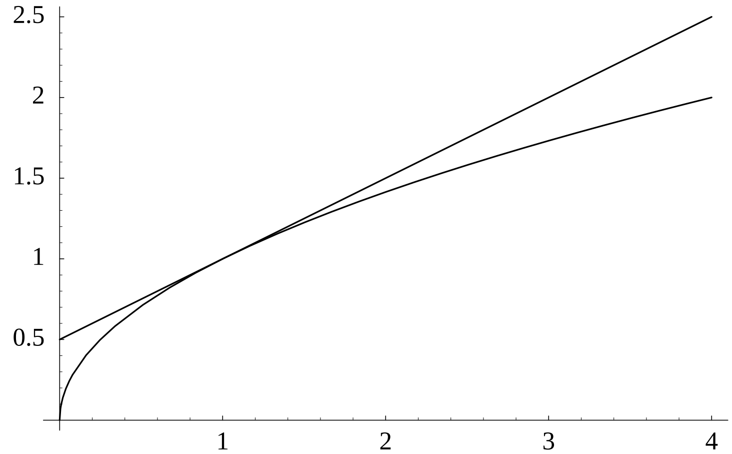

Figure 3.1.4 Graph of $f(x) = \sqrt{x}$ and its tangent line at $(1, 1)$

will call the graph of T, which is a straight line through $(1, 1)$ with slope 0.5, the *tangent line* to the graph of f at $(1, 1)$. See Figure 3.1.4.

As an example of using T to approximate f, note that, to 4 decimal places,

$$\sqrt{1.1} = 1.0488,$$

while

$$T(1.1) = \frac{1}{2}(1.1 - 1) + 1 = 1.05,$$

giving a remainder of only

$$R(0.1) = 1.0488 - 1.0500 = -0.0012.$$

This approximation is remarkably accurate considering the simplicity of the calculations used to obtain it. Of course, we expect the accuracy of the approximation to increase as h decreases. For example, to 4 decimal places,

$$\sqrt{1.05} = 1.0247,$$

while

$$T(1.05) = \frac{1}{2}(1.05 - 1) + 1 = 1.025,$$

giving a remainder of only

$$R(0.05) = 1.0247 - 1.0250 = -0.0003.$$

Note that when we decreased h from 0.1 to 0.05, a factor of $\frac{1}{2}$, the error went from -0.0012 to -0.003, a factor of $\frac{1}{4}$. This is evidence of the quadratic nature of the error, the fact that $R(h)$ is approaching 0 like h^2, not like h.

Using the ideas of this example, we may now make the following definition.

Definition Let f be a function defined in an open interval about a point c. If T is an affine function such that $T(c) = f(c)$ and

$$R(h) = f(c+h) - T(c+h)$$

is $o(h)$, then we call T the *best affine approximation* to f at c. Moreover, the graph of T is called the *tangent line* to the graph of f at $(c, f(c))$.

Using the point-slope form for the equation of a line, the equation of the tangent line at $(c, f(c))$ may be written in the form

$$y - f(c) = m(x - c)$$

for some slope m. That is,

$$y = m(x - c) + f(c),$$

or, in other words, the best affine approximation has the form

$$T(x) = m(x - c) + f(c).$$

Thus, to determine T, we need only find the value of m. Since this number m is of such importance, we give it a formal definition.

Definition If

$$T(x) = m(x - c) + f(c)$$

is the best affine approximation to f at c, then we call m, the slope of the graph of T, the *derivative* of f at c. This value is denoted by $f'(c)$.

With this notation, the best affine approximation has the form

$$T(x) = f'(c)(x - c) + f(c). \tag{3.1.8}$$

Example As a consequence of our previous example, if $f(x) = \sqrt{x}$, then

$$f'(1) = \frac{1}{2}.$$

Example Let $f(x) = x^2$ and suppose we wish to find the best affine approximation to f at 3. Then $f(3) = 9$, so we will let

$$T(x) = m(x - 3) + 9$$

and

$$R(h) = f(3+h) - T(3+h) = (3+h)^2 - (mh + 9).$$

Hence

$$R(h) = 9 + 6h + h^2 - mh - 9 = h(6 + h - m),$$

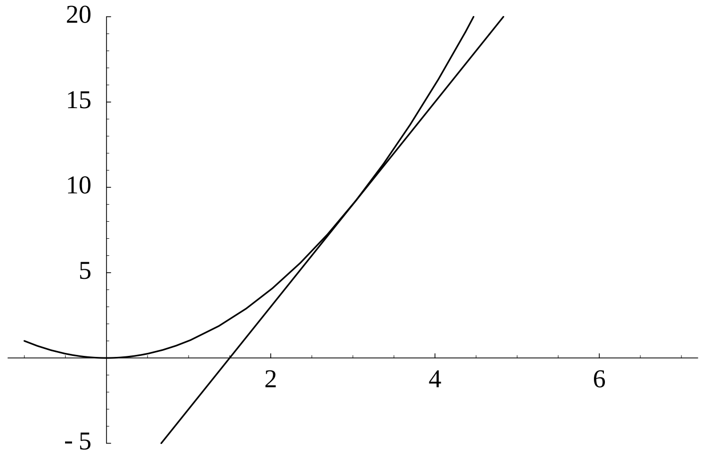

Figure 3.1.5 Graph of $f(x) = x^2$ and its tangent line at $(3, 9)$

and so

$$\lim_{h \to 0} \frac{R(h)}{h} = \lim_{h \to 0} \frac{h(6 + h - m)}{h} = \lim_{h \to 0} (6 - m + h) = 6 - m.$$

Thus $R(h)$ is $o(h)$ if and only if $m = 6$. It follows then that $f'(3) = 6$ and the best affine approximation to f at 3 is

$$T(x) = 6(x - 3) + 9.$$

The equation of the tangent line at $(3, 9)$ is

$$y = 6(x - 3) + 9,$$

or, equivalently,

$$y = 6x - 9.$$

See Figure 3.1.5.

In Sections 3.2 through 3.5 we will explore techniques which will simplify greatly the process of finding derivatives.

Problems

1. Consider the problem of finding an affine approximation for $f(x) = \sin(x)$ near 0. Since $f(0) = 0$, we let $T(x) = mx$ and

$$R(h) = f(h) - T(h) = \sin(h) - mh.$$

   (a) Plot $|R(h)|$ on the interval $[-0.2, 0.2]$ for $m = 0.0, 0.2, 0.4, \ldots, 2.0$.

   (b) Which value of m gives the smallest errors?

2. For each of the following, decide if the given function is $O(h)$, $o(h)$, or neither.

   (a) $f(x) = x^3$ (b) $f(x) = x^2 + 3x$

(c) $g(t) = 4t^3 - 3t^2$

(d) $g(x) = \sqrt{4+x} - \dfrac{x}{4} - 2$

(e) $f(t) = t^{\frac{4}{3}}$

(f) $g(t) = t - t^{\frac{3}{5}}$

3. Let $f(x) = \sqrt{x}$, $T(x) = \dfrac{1}{2}(x-9)+3$, and $S(x) = \dfrac{1}{6}(x-9)+3$.
   (a) Graph f, T, and S together. Note that the graphs of T and S are straight lines passing through the point $(9,3)$ on the graph of f.
   (b) Let $R_T(h) = f(9+h) - T(9+h)$. Is $R_T(h)$ $o(h)$? Is it $O(h)$?
   (c) Let $R_S(h) = f(9+h) - S(9+h)$. Is $R_S(h)$ $o(h)$? Is it $O(h)$?
   (d) Which of T or S is the best affine approximation to f at 9?
   (e) Use the best affine approximation to f at 9 to approximate $\sqrt{10}$, $\sqrt{8.9}$, and $\sqrt{9.3}$. Compare these approximations with values from your calculator.

4. Let $g(z) = z^2$, $T(z) = 2(z-1)+1$, and $S(z) = 3(z-1)+1$.
   (a) Graph g, T, and S together. Note that the graphs of T and S are straight lines passing through the point $(1,1)$ on the graph of g.
   (b) Let $R_T(h) = g(1+h) - T(1+h)$. Is $R_T(h)$ $o(h)$? Is it $O(h)$?
   (c) Let $R_S(h) = g(1+h) - S(1+h)$. Is $R_S(h)$ $o(h)$? Is it $O(h)$?
   (d) Which of T or S is the best affine approximation to g at 1?
   (e) Use the best affine approximation to g at 1 to approximate $(1.1)^2$ and $(0.999)^2$. Compare these approximations with values from your calculator.

5. Find the best affine approximation to $f(x) = 2x^2$ at 1. What is $f'(1)$?

6. Find the best affine approximation to $g(x) = \dfrac{1}{x}$ at 1. What is $g'(1)$?

7. Find the best affine approximation to $f(t) = t^2 + t - 1$ at 0. What is $f'(0)$?