Approximating Functions Near a Specified Point

Suppose that you are interested in the values of some function $f(x)$ for x near some fixed point x_0. The function is too complicated to work with directly. So you wish to work instead with some other function $F(x)$ that is both simple and a good approximation to $f(x)$ for x near x_0. We'll consider a couple of examples of this scenario later. First, we develop several different approximations.

First approximation

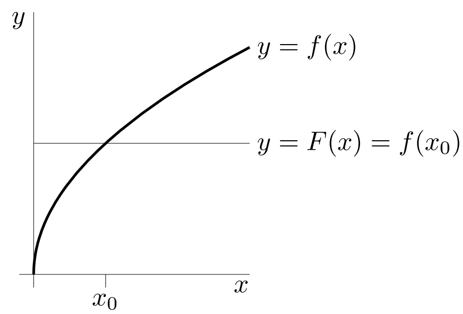
The simplest functions are those that are constants. The first approximation will be by a constant function. That is, the approximating function will have the form $F(x) = A$. To ensure that $F(x)$ is a good approximation for x close to x_0, we chose the constant A so that $f(x)$ and $F(x)$ take exactly the same value when $x = x_0$.

$$F(x) = A \qquad \text{so} \qquad F(x_0) = A = f(x_0) \implies A = f(x_0)$$

Our first, and crudest, approximation rule is

$$f(x) \approx f(x_0) \tag{1}$$

Here is a figure showing the graphs of a typical $f(x)$ and approximating function $F(x)$. At $x = x_0$, $f(x)$



and $F(x)$ take the same value. For x very near x_0, the values of $f(x)$ and $F(x)$ remain close together. But the quality of the approximation deteriorates fairly quickly as x moves away from x_0.

Second Approximation – the tangent line, or linear, approximation

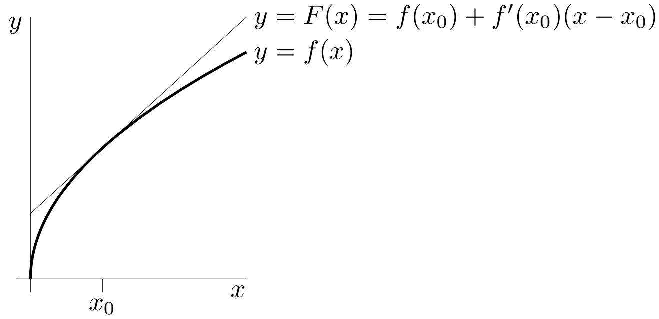
We now develop a better approximation by allowing the approximating function to be a linear function of x and not just a constant function. That is, we allow $F(x)$ to be of the form $A + Bx$. To ensure that $F(x)$ is a good approximation for x close to x_0, we chose the constants A and B so that $f(x_0) = F(x_0)$ and $f'(x_0) = F'(x_0)$. Then $f(x)$ and $F(x)$ will have both the same value and the same slope at $x = x_0$.

$$\begin{aligned} F(x) &= A + Bx & &\implies & F(x_0) &= A + Bx_0 = f(x_0) \\ F'(x) &= B & &\implies & F'(x_0) &= \qquad\quad B = f'(x_0) \end{aligned}$$

Subbing $B = f'(x_0)$ into $A + Bx_0 = f(x_0)$ gives $A = f(x_0) - x_0 f'(x_0)$ and consequently $F(x) = A + Bx = f(x_0) - x_0 f'(x_0) + x f'(x_0) = f(x_0) + f'(x_0)(x - x_0)$. So, our second approximation is

$$\boxed{f(x) \approx f(x_0) + f'(x_0)(x - x_0)} \tag{2}$$

Here is a figure showing the graphs of a typical $f(x)$ and approximating function $F(x)$. Observe that the



graph of $f(x_0) + f'(x_0)(x - x_0)$ remains close to the graph of $f(x)$ for a much larger range of x than did the graph of $f(x_0)$.

Third approximation – the quadratic approximation

We finally develop a still better approximation by allowing the approximating function be to a quadratic function of x. That is, we allow $F(x)$ to be of the form $A + Bx + Cx^2$. To ensure that $F(x)$ is a good approximation for x close to x_0, we chose the constants A, B and C so that $f(x_0) = F(x_0)$ and $f'(x_0) = F'(x_0)$ and $f''(x_0) = F''(x_0)$.

$$\begin{aligned}
F(x) &= A + Bx + Cx^2 &\implies\quad F(x_0) &= A + Bx_0 + Cx_0^2 = f(x_0)\\
F'(x) &= B + 2Cx &\implies\quad F'(x_0) &= B + 2Cx_0 = f'(x_0)\\
F''(x) &= 2C &\implies\quad F''(x_0) &= 2C = f''(x_0)
\end{aligned}$$

Solve for C first, then B and finally A.

$$\begin{aligned}
C = \tfrac{1}{2}f''(x_0) &\implies B = f'(x_0) - 2Cx_0 = f'(x_0) - x_0 f''(x_0)\\
&\implies A = f(x_0) - x_0 B - Cx_0^2 = f(x_0) - x_0[f'(x_0) - x_0 f''(x_0)] - \tfrac{1}{2}f''(x_0)x_0^2
\end{aligned}$$

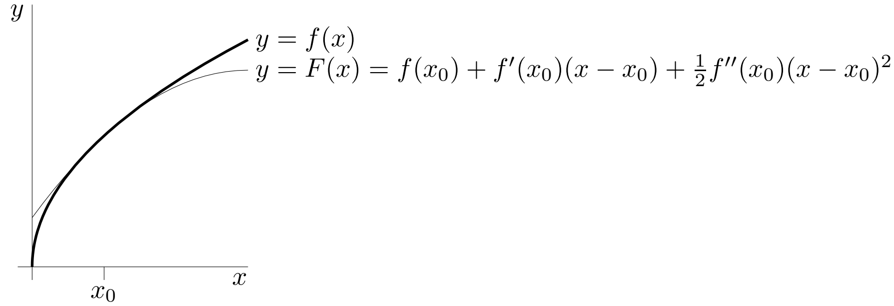
Then build up $F(x)$.

$$\begin{aligned}
F(x) &= f(x_0) - f'(x_0)x_0 + \tfrac{1}{2}f''(x_0)x_0^2 && \text{(this line is } A)\\
&\quad + f'(x_0)\,x - f''(x_0)x_0 x && \text{(this line is } Bx)\\
&\quad + \tfrac{1}{2}f''(x_0)x^2 && \text{(this line is } Cx^2)\\
&= f(x_0) + f'(x_0)(x - x_0) + \tfrac{1}{2}f''(x_0)(x - x_0)^2
\end{aligned}$$

Our third approximation is

$$\boxed{f(x) \approx f(x_0) + f'(x_0)(x - x_0) + \tfrac{1}{2}f''(x_0)(x - x_0)^2} \tag{3}$$

It is called the quadratic approximation. Here is a figure showing the graphs of a typical $f(x)$ and approximating function $F(x)$. The third approximation looks better than both the first and second.



Still Better Approximations – Taylor Polynomials

We can use the same strategy to generate still better approximations by polynomials of any degree we like. Let's approximate by a polynomial of degree n. The algebra will be simpler if we make the approximating polynomial $F(x)$ of the form

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

Because x_0 is itself a constant, this is really just a rewriting of $A_0 + A_1x + A_2x^2 + \cdots + A_nx^n$. For example,

$$\begin{aligned} a_0 + a_1(x-x_0) + a_2(x-x_0)^2 &= a_0 + a_1x - a_1x_0 + a_2x^2 - 2a_2xx_0 + a_2x_0^2 \\ &= (a_0 - a_1x_0 + a_2x_0^2) + (a_1 - 2a_2x_0)x + a_2x^2 \\ &= A_0 + A_1x + A_2x^2 \end{aligned}$$

with $A_0 = a_0 - a_1x_0 + a_2x_0^2$, $A_1 = a_1 - 2a_2x_0$ and $A_2 = a_2$. The advantage of the form $a_0 + a_1(x-x_0) + \cdots$ is that $x - x_0$ is zero when $x = x_0$, so lots of terms in the computation drop out. We determine the coefficients a_i by the requirements that $f(x)$ and its approximator $F(x)$ have the same value and the same first n derivatives at $x = x_0$.

$$\begin{aligned}
F(x) &= a_0 + a_1(x-x_0) + a_2(x-x_0)^2 + \cdots + a_n(x-x_0)^n & \implies && F(x_0) = a_0 = f(x_0) \\
F'(x) &= a_1 + 2a_2(x-x_0) + 3a_3(x-x_0)^2 + \cdots + na_n(x-x_0)^{n-1} & \implies && F'(x_0) = a_1 = f'(x_0) \\
F''(x) &= 2a_2 + 3\times 2a_3(x-x_0) + \cdots + n(n-1)a_n(x-x_0)^{n-2} & \implies && F''(x_0) = 2a_2 = f''(x_0) \\
F^{(3)}(x) &= 3\times 2a_3 + \cdots + n(n-1)(n-2)a_n(x-x_0)^{n-3} & \implies && F^{(3)}(x_0) = 3\times 2a_3 = f^{(3)}(x_0) \\
&\vdots && & \vdots \\
F^{(n)}(x) &= n!a_n & \implies && F^{(n)}(x_0) = n!a_n = f^{(n)}(x_0)
\end{aligned}$$

Here $n! = n(n-1)(n-2)\cdots 1$ is called n factorial. Hence

$$a_0 = f(x_0) \quad a_1 = f'(x_0) \quad a_2 = \tfrac{1}{2!}f''(x_0) \quad a_3 = \tfrac{1}{3!}f^{(3)}(x_0) \quad \cdots \quad a_n = \tfrac{1}{n!}f^{(n)}(x_0)$$

and the approximator, which is called the Taylor polynomial of degree n for $f(x)$ at $x = x_0$, is

$$\boxed{f(x) \approx f(x_0) + f'(x_0)\,(x-x_0) + \tfrac{1}{2!}f''(x_0)\,(x-x_0)^2 + \tfrac{1}{3!}f^{(3)}(x_0)\,(x-x_0)^3 + \cdots + \tfrac{1}{n!}f^{(n)}(x_0)\,(x-x_0)^n} \tag{4}$$

or, in summation notation,

$$\boxed{f(x) \approx \sum_{\ell=0}^{n} \tfrac{1}{\ell!}f^{(n)}(x_0)\,(x-x_0)^\ell} \tag{4}$$

where we are using the standard convention that $0! = 1$.

Another Notation

Suppose that we have two variables x and y that are related by $y = f(x)$, for some function f. For example, x might be the number of cars manufactured per week in some factory and y the cost of manufacturing those x cars. Let x_0 be some fixed value of x and let $y_0 = f(x_0)$ be the corresponding value of y. Now suppose that x changes by an amount Δx, from x_0 to $x_0 + \Delta x$. As x undergoes this change, y changes from $y_0 = f(x_0)$ to $f(x_0 + \Delta x)$. The change in y that results from the change Δx in x is

$$\Delta y = f(x_0 + \Delta x) - f(x_0)$$

Substituting $x = x_0 + \Delta x$ into the linear approximation (2) yields the approximation

$$f(x_0 + \Delta x) \approx f(x_0) + f'(x_0)(x_0 + \Delta x - x_0) = f(x_0) + f'(x_0)\Delta x$$

for $f(x_0 + \Delta x)$ and consequently the approximation

$$\Delta y = f(x_0 + \Delta x) - f(x_0) \approx f(x_0) + f'(x_0)\Delta x - f(x_0) \implies \boxed{\Delta y \approx f'(x_0)\Delta x} \tag{5}$$

for Δy. In the automobile manufacturing example, when the production level is x_0 cars per week, increasing the production level by Δx will cost approximately $f'(x_0)\Delta x$. The additional cost per additional car, $f'(x_0)$, is called the "marginal cost" of a car.

If we use the quadratic approximation (3) in place of the linear approximation (2)

$$f(x_0 + \Delta x) \approx f(x_0) + f'(x_0)\Delta x + \tfrac{1}{2} f''(x_0)\Delta x^2$$

we arrive at the quadratic approximation

$$\Delta y = f(x_0 + \Delta x) - f(x_0) \approx f(x_0) + f'(x_0)\Delta x + \tfrac{1}{2} f''(x_0)\Delta x^2 - f(x_0) \implies \boxed{\Delta y \approx f'(x_0)\Delta x + \tfrac{1}{2} f''(x_0)\Delta x^2} \tag{6}$$

for Δy.

Example 1

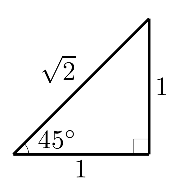
Suppose that you wish to compute, approximately, $\tan 46°$, but that you can't just use your calculator. This will be the case, for example, if the computation is an exercise to help prepare you for designing the software to be used by the calculator.

In this example, we choose $f(x) = \tan x$, $x = 46\frac{\pi}{180}$ radians and $x_0 = 45\frac{\pi}{180} = \frac{\pi}{4}$ radians. This is a good choice for x_0 because

- $x_0 = 45°$ is close to $x = 46°$. Generally, the closer x is to x_0, the better the quality of our various approximations
- We know the values of all trig functions at $45°$.

The first step in applying our approximations is to compute f and its first two derivatives at $x = x_0$.

$$\begin{aligned} f(x) &= \tan x & &\implies & f(x_0) &= \tan\tfrac{\pi}{4} = 1 \\ f'(x) &= (\cos x)^{-2} & &\implies & f'(x_0) &= \tfrac{1}{\cos^2(\pi/4)} = 2 \\ f''(x) &= -2\tfrac{-\sin x}{\cos^3 x} & &\implies & f''(x_0) &= 2\tfrac{\sin(\pi/4)}{\cos^3(\pi/4)} = 2\tfrac{1/\sqrt{2}}{(1/\sqrt{2})^3} = 2\tfrac{1}{1/2} = 4 \end{aligned}$$



As $x - x_0 = 46\frac{\pi}{180} - 45\frac{\pi}{180} = \frac{\pi}{180}$ radians, the three approximations are

$$\begin{aligned} f(x) &\approx f(x_0) &&= 1 \\ f(x) &\approx f(x_0) + f'(x_0)(x - x_0) &&= 1 + 2\tfrac{\pi}{180} = 1.034907 \\ f(x) &\approx f(x_0) + f'(x_0)(x - x_0) + \tfrac{1}{2} f''(x_0)(x - x_0)^2 &&= 1 + 2\tfrac{\pi}{180} + \tfrac{1}{2}4\big(\tfrac{\pi}{180}\big)^2 = 1.035516 \end{aligned}$$

For comparison purposes, $\tan 46°$ really is 1.035530 to 6 decimal places.

Recall that all of our derivative formulae for trig functions, were developed under the assumption that angles were measured in radians. As our approximation formulae used those derivatives, we were obliged to express $x - x_0$ in radians.

Example 2

Let's find all Taylor polyomial for $\sin x$ and $\cos x$ at $x_0 = 0$. To do so we merely need compute all derivatives of $\sin x$ and $\cos x$ at $x_0 = 0$. First, compute all derivatives at general x.

$$\begin{array}{llllll} f(x) = \sin x & f'(x) = \cos x & f''(x) = -\sin x & f^{(3)}(x) = -\cos x & f^{(4)}(x) = \sin x & \cdots \\ g(x) = \cos x & g'(x) = -\sin x & g''(x) = -\cos x & g^{(3)}(x) = \sin x & g^{(4)}(x) = \cos x & \cdots \end{array}$$

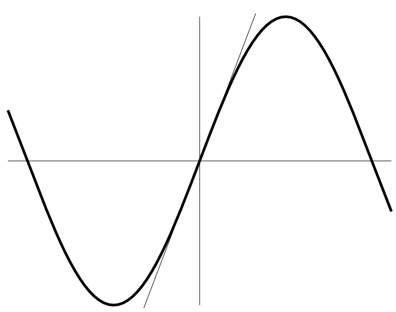
The pattern starts over again with the fourth derivative being the same as the original function. Now set $x = x_0 = 0$.

$$\begin{array}{lllllll} f(x) = \sin x & f(0) = 0 & f'(0) = 1 & f''(0) = 0 & f^{(3)}(0) = -1 & f^{(4)}(0) = 0 & \cdots \\ g(x) = \cos x & g(0) = 1 & g'(0) = 0 & g''(0) = -1 & g^{(3)}(0) = 0 & g^{(4)}(0) = 1 & \cdots \end{array}$$

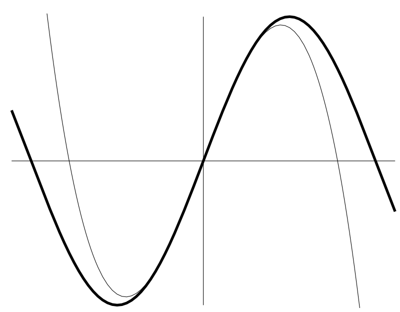
For $\sin x$, all even numbered derivatives are zero. The odd numbered derivatives alternate between 1 and -1. For $\cos x$, all odd numbered derivatives are zero. The even numbered derivatives alternate between 1 and -1. So, the Taylor polynomials that best approximate $\sin x$ and $\cos x$ near $x = x_0 = 0$ are

$$\begin{aligned} \sin x &\approx x - \tfrac{1}{3!}x^3 + \tfrac{1}{5!}x^5 - \cdots \\ \cos x &\approx 1 - \tfrac{1}{2!}x^2 + \tfrac{1}{4!}x^4 - \cdots \end{aligned}$$

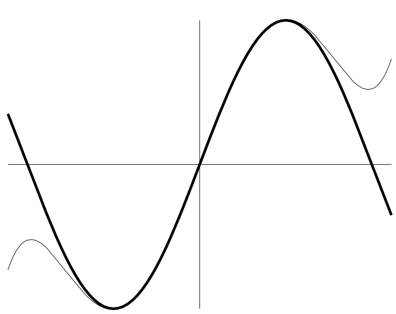
Here are graphs of $\sin x$ and its Taylor poynomials (about $x_0 = 0$) up to degree seven.



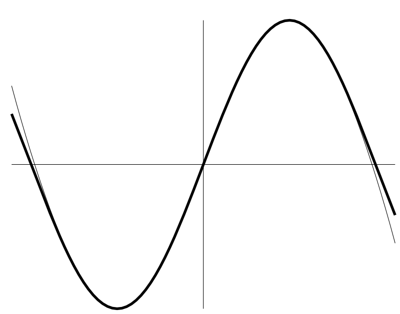
$\sin x \approx x$



$\sin x \approx x - \frac{1}{3!}x^3$



$\sin x \approx x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5$



$\sin x \approx x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \frac{1}{7!}x^7$

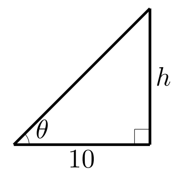
To get an idea of how good these Taylor polynomials are at approximating sin and cos, let's concentrate on $\sin x$ and consider x's whose magnitude $|x| \le 1$. (If you're writing software to evaluate $\sin x$, you can always use the trig identity $\sin(x) = \sin(x - 2n\pi)$, to easily restrict to $|x| \le \pi$, and then use the trig identity $\sin(x) = -\sin(x \pm \pi)$ to reduce to $|x| \le \frac{\pi}{2}$ and then use the trig identity $\sin(x) = \cos(\frac{\pi}{2} \pm x))$ to reduce to

$|x| \le \frac{\pi}{4}$.) If $|x| \le 1$ radians (recall that the derivative formulae that we used to derive the Taylor polynomials are valid only when x is in radians), or equivalently if $|x|$ is no larger than $\frac{180}{\pi} \approx 57°$, then the magnitudes of the successive terms in the Taylor polynomials for $\sin x$ are bounded by

$$\begin{array}{lll} |x| \le 1 & \frac{1}{3!}|x|^3 \le \frac{1}{6} & \frac{1}{5!}|x|^3 \le \frac{1}{120} \approx 0.0083 \\ \frac{1}{7!}|x|^7 \le \frac{1}{7!} \approx 0.0002 & \frac{1}{9!}|x|^9 \le \frac{1}{9!} \approx 0.000003 & \frac{1}{11!}|x|^{11} \le \frac{1}{11!} \approx 0.000000025 \end{array}$$

From these inequalities, and the graphs on the previous page, it certainly looks like, for x not too large, even relatively low degree Taylor polynomials give very good approximations. We'll see later how to get rigorous error bounds on our Taylor polynomial approximations.

Example 3



Suppose that you are ten meters from a vertical pole. You were contracted to measure the height of the pole. You can't take it down or climb it. So you measure the angle subtended by the top of the pole. You measure $\theta = 30°$, which gives $h = 10 \tan 30° = \frac{10}{\sqrt{3}} \approx 5.77$m. But there's a catch. Angles are hard to measure accurately. Your contract specifies that the height must be measured to within an accuracy of 10 cm. How accurate did your measurement of θ have to be?

For simplicity, we are going to assume that the pole is perfectly straight and perfectly vertical and that your distance from the pole was exactly 10 m. Write $h = h_0 + \Delta h$, where h is the exact height and $h_0 = \frac{10}{\sqrt{3}}$ is the computed height. Their difference, Δh, is the error. Similarly, write $\theta = \theta_0 + \Delta\theta$ where θ is the exact angle, θ_0 is the measured angle and $\Delta\theta$ is the error. Then

$$h_0 = 10 \tan \theta_0 \qquad h_0 + \Delta h = 10 \tan(\theta_0 + \Delta\theta)$$

We apply $\Delta y \approx f'(x_0)\Delta x$, with y replaced by h and x replaced by θ. That is, we apply $\Delta h \approx f'(\theta_0)\Delta\theta$. Choosing $f(\theta) = 10 \tan \theta$ and $\theta_0 = 30°$ and subbing in $f'(\theta_0) = 10 \sec^2 \theta_0 = 10 \sec^2 30° = 10\big(\frac{2}{\sqrt{3}}\big)^2 = \frac{40}{3}$, we see that the error in the computed value of h and the error in the measured value of θ are related by

$$\Delta h \approx \tfrac{40}{3}\Delta\theta$$

To achieve $|\Delta h| \le .1$, we better have $|\Delta\theta|$ smaller than $.1\frac{3}{40}$ radians or $.1\frac{3}{40}\frac{180}{\pi} = .43°$.

Example 4

The radius of a sphere is measured with a percentage error of at most ε%. Find the approximate percentage error in the surface area and volume of the sphere.

Solution. Suppose that the exact radius is r_0 and that the measured radius is $r_0 + \Delta r$. Then the absolute error in the measurement is $|\Delta r|$ and the percentage error is $100\frac{|\Delta r|}{r_0}$. We are told that $100\frac{|\Delta r|}{r_0} \le \varepsilon$. The surface area of a sphere of radius r is $A(r) = 4\pi r^2$. The error in the surface area computed with the measured radius is

$$\Delta A = A(r_0 + \Delta r) - A(r_0) \approx A'(r_0)\Delta r$$

The corresponding percentage error is

$$100\frac{|\Delta A|}{A(r_0)} \approx 100\frac{|A'(r_0)\Delta r|}{A(r_0)} = 100\frac{8\pi r_0|\Delta r|}{4\pi r_0^2} = 2 \times 100\frac{|\Delta r|}{r_0} \le \boxed{2\varepsilon}$$

The volume of a sphere of radius r is $V(r) = \frac{4}{3}\pi r^3$. The error in the volume computed with the measured radius is

$$\Delta V = V(r_0 + \Delta r) - V(r_0) \approx V'(r_0)\Delta r$$

The corresponding percentage error is

$$100\frac{|\Delta V|}{V(r_0)} \approx 100\frac{|V'(r_0)\Delta r|}{V(r_0)} = 100\frac{4\pi r_0^2|\Delta r|}{4\pi r_0^3/3} = 3 \times 100\frac{|\Delta r|}{r_0} \le \boxed{3\varepsilon}$$

We have just computed an approximation to ΔV. In this problem, we can compute the exact error

$$V(r_0 + \Delta r) - V(r_0) = \tfrac{4}{3}\pi(r_0 + \Delta r)^3 - \tfrac{4}{3}\pi r_0^3$$

Applying $(a + b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$ with $a = r_0$ and $b = \Delta r$, gives

$$\begin{aligned} V(r_0 + \Delta r) - V(r_0) &= \tfrac{4}{3}\pi[r_0^3 + 3r_0^2\Delta r + 3r_0\,\Delta r^2 + \Delta r^3 - r_0^3] \\ &= \tfrac{4}{3}\pi[3r_0^2\Delta r + 3r_0\,\Delta r^2 + \Delta r^3] \end{aligned}$$

The linear approximation, $\Delta V \approx 4\pi r_0^2 \times \Delta r$, is recovered by retaining only the first of the three terms in the square brackets. Thus the difference between the exact error and the linear approximation to the error is obtained by retaining only the last two terms in the square brackets. This has magnitude

$$\tfrac{4}{3}\pi\big|3r_0\,\Delta r^2 + \Delta r^3\big| = \tfrac{4}{3}\pi\big|3r_0 + \Delta r\big|\Delta r^2$$

or in percentage terms

$$100\frac{1}{\frac{4}{3}\pi r_0^3}\tfrac{4}{3}\pi\big|3r_0\,\Delta r^2 + \Delta r^3\big| = 100\big|3\tfrac{\Delta r^2}{r_0^2} + \tfrac{\Delta r^3}{r_0^3}\big| = \big(100\tfrac{3\Delta r}{r_0}\big)\big(\tfrac{\Delta r}{r_0}\big)\big|1 + \tfrac{\Delta r}{3r_0}\big| \le 3\varepsilon\big(\tfrac{\varepsilon}{100}\big)\big(1 + \tfrac{\varepsilon}{300}\big)$$

Thus the difference between the exact error and the linear approximation is roughly a factor of $\frac{\varepsilon}{100}$ smaller than the linear approximation 3ε.

Example 5

If an aircraft crosses the Atlantic ocean at a speed of u mph, the flight costs the company

$$C(u) = 100 + \tfrac{u}{3} + \tfrac{240{,}000}{u}$$

dollars per passenger. When there is no wind, the aircraft flies at an airspeed of 550mph. Find the approximate savings, per passenger, when there is a 35 mph tail wind. Estimate the cost when there is a 50 mph head wind.

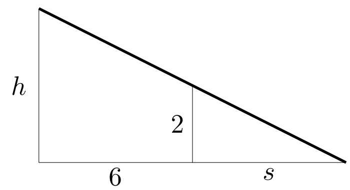
Solution. Let $u_0 = 550$. When the aircraft flies at speed u_0, the cost per passenger is $C(u_0)$. By (5), a change of Δu in the airspeed results in an change of

$$\Delta C \approx C'(u_0)\Delta u = \big[\tfrac{1}{3} - \tfrac{240{,}000}{u_0^2}\big]\Delta u = \big[\tfrac{1}{3} - \tfrac{240{,}000}{550^2}\big]\Delta u \approx -.460\Delta u$$

in the cost per passenger. With the tail wind $\Delta u = 35$ and the resulting $\Delta C \approx -.460 \times 35 = -16.10$, so there is a savings of $\boxed{\$16.10}$. With the head wind $\Delta u = -50$ and the resulting $\Delta C \approx -.4601 \times (-50) = 23.01$, so there is an additional cost of $\boxed{\$23.00}$.

Example 6

To compute the height h of a lamp post, the length s of the shadow of a two meter pole is measured. The pole is 6 m from the lamp post. If the length of the shadow was measured to be 4 m, with an error of at most one cm, find the height of the lamp post and estimate the relative error in the height.



Solution. By similar triangles,

$$\frac{s}{2} = \frac{6+s}{h} \implies h = (6+s)\frac{2}{s} = \frac{12}{s} + 2$$

The length of the shadow was measured to be $s_0 = 4$ m. The corresponding height of the lamp post is $h_0 = \frac{12}{s_0} + 2 = \frac{12}{4} + 2 = \boxed{5 \text{ m}}$. If the error in the measurement of the length of the shadow was Δs, then the exact shadow length was $s = s_0 + \Delta s$ and the exact lamp post height is $h = f(s_0 + \Delta s)$, where $f(s) = \frac{12}{s} + 2$. The error in the computed lamp post height is $\Delta h = h - h_0 = f(s_0 + \Delta s) - f(s_0)$. By (5)

$$\Delta h \approx f'(s_0)\Delta s = -\tfrac{12}{s_0^2}\Delta s = -\tfrac{12}{4^2}\Delta s$$

We are told that $|\Delta s| \le \frac{1}{10}$ m. Consequently $|\Delta h| \le \frac{12}{4^2}\frac{1}{10} = \frac{3}{40}$ (approximately). The relative error is then approximately

$$\tfrac{|\Delta h|}{h_0} \le \tfrac{3}{40\times 5} = \boxed{0.015} \text{ or } \boxed{1.5\%}$$

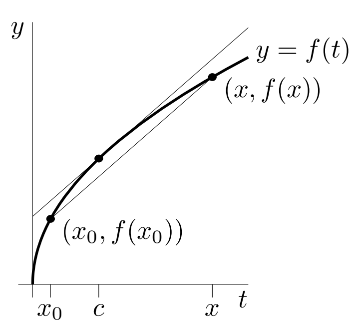
The Error in the Approximations

Any time you make an approximation, it is desirable to have some idea of the size of the error you introduced. We will now develop a formula for the error introduced by the approximation $f(x) \approx f(x_0)$. This formula can be used to get an upper bound on the size of the error, even when you cannot determine $f(x)$ exactly.

By simple algebra

$$f(x) = f(x_0) + \frac{f(x) - f(x_0)}{x - x_0}(x - x_0) \tag{7}$$

The coefficient $\frac{f(x)-f(x_0)}{x-x_0}$ of $(x - x_0)$ is the average slope of $f(t)$ as t moves from $t = x_0$ to $t = x$. In the figure below, it is the slope of the secant joining the points $(x_0, f(x_0))$ and $(x, f(x))$. As t moves x_0 to



x, the instantaneous slope $f'(t)$ keeps changing. Sometimes it is larger than the average slope $\frac{f(x)-f(x_0)}{x-x_0}$

and sometimes it is smaller than the average slope. But, by the Mean–Value Theorem, there must be some number c, strictly between x_0 and x, for which $f'(c) = \frac{f(x)-f(x_0)}{x-x_0}$. Subbing this into formula (7)

$$\boxed{f(x) = f(x_0) + f'(c)(x - x_0) \qquad \text{for some } c \text{ strictly between } x_0 \text{ and } x} \tag{8}$$

Thus the error in the approximation $f(x) \approx f(x_0)$ is exactly $f'(c)(x - x_0)$ for some c strictly between x_0 and x. There are formulae similar to (8), that can be used to bound the error in our other approximations. One is

$$\boxed{f(x) = f(x_0) + f'(x_0)(x - x_0) + \tfrac{1}{2} f''(c)(x - x_0)^2 \qquad \text{for some } c \text{ strictly between } x_0 \text{ and } x}$$

It implies that the error in the approximation $f(x) \approx f(x_0) + f'(x_0)\,(x - x_0)$ is exactly $\frac{1}{2} f''(c)\,(x - x_0)^2$ for some c strictly between x_0 and x. In general

$$\begin{aligned} f(x) =& f(x_0) + f'(x_0)\,(x - x_0) + \cdots + \tfrac{1}{n!} f^{(n)}(x_0)\,(x - x_0)^n \\ &+ \tfrac{1}{(n+1)!} f^{(n+1)}(c)\,(x - x_0)^{n+1} \qquad \text{for some } c \text{ strictly between } x_0 \text{ and } x \end{aligned}$$

That is, the error introduced when $f(x)$ is approximated by its Taylor polynomial of degree n, is precisely the last term of the Taylor polynomial of degree $n + 1$, but with the derivative evaluated at some point between x_0 and x, rather than exactly at x_0. These error formulae are proven in a supplement (which you are not responsible for) at the end of these notes.

Example 7

Suppose we wish to approximate $\sin 46°$ using Taylor polynomials about $x_0 = 45°$. Then, we would define

$$f(x) = \sin x \qquad x_0 = 45° = 45\tfrac{\pi}{180}\text{radians} \qquad x = 46° = 46\tfrac{\pi}{180}\text{radians} \qquad x - x_0 = \tfrac{\pi}{180}\text{radians}$$

The first few derivatives of f at x_0 are

$$\begin{aligned} f(x) &= \sin x & f(x_0) &= \tfrac{1}{\sqrt{2}} \\ f'(x) &= \cos x & f'(x_0) &= \tfrac{1}{\sqrt{2}} \\ f''(x) &= -\sin x & f''(x_0) &= -\tfrac{1}{\sqrt{2}} \\ f^{(3)}(x) &= -\cos x \end{aligned}$$

The constant, linear and quadratic approximations for $\sin 46°$

$$\begin{aligned} \sin 46° &\approx f(x_0) &&= \tfrac{1}{\sqrt{2}} &&= 0.70710678 \\ \sin 46° &\approx f(x_0) + f'(x_0)(x - x_0) &&= \tfrac{1}{\sqrt{2}} + \tfrac{1}{\sqrt{2}}\left(\tfrac{\pi}{180}\right) &&= 0.71944812 \\ \sin 46° &\approx f(x_0) + f'(x_0)(x - x_0) + \tfrac{1}{2} f''(x_0)(x - x_0)^2 &&= \tfrac{1}{\sqrt{2}} + \tfrac{1}{\sqrt{2}}\left(\tfrac{\pi}{180}\right) - \tfrac{1}{\sqrt{2}}\left(\tfrac{\pi}{180}\right)^2 &&= 0.71934042 \end{aligned}$$

The corresonding errors are

$$\begin{aligned} \text{error in } 0.70710678 &= f'(c)(x - x_0) &&= \cos c\,\left(\tfrac{\pi}{180}\right) \\ \text{error in } 0.71944812 &= \tfrac{1}{2} f''(c)(x - x_0)^2 &&= -\tfrac{1}{2}\sin c\,\left(\tfrac{\pi}{180}\right)^2 \\ \text{error in } 0.71923272 &= \tfrac{1}{3!} f'(c)(x - x_0)^3 &&= -\tfrac{1}{3!}\cos c\,\left(\tfrac{\pi}{180}\right)^3 \end{aligned}$$

In each of these three cases c must lie somewhere between 45° and 46°. No matter what c is, we know that $|\sin c| \le 1$ and $\cos c| \le 1$. Hence

$$\begin{aligned}
\big|\text{error in } 0.70710678\big| &\le \big(\tfrac{\pi}{180}\big) &&< 0.018\\
\big|\text{error in } 0.71944812\big| &\le \tfrac{1}{2}\big(\tfrac{\pi}{180}\big)^2 &&< 0.00015\\
\big|\text{error in } 0.71934042\big| &\le \tfrac{1}{3!}\big(\tfrac{\pi}{180}\big)^3 &&< 0.0000009
\end{aligned}$$

Example 2 Revisited

In the second example (measuring the height of the pole), we used the linear approximation

$$f(\theta_0 + \Delta\theta) \approx f(\theta_0) + f'(\theta_0)\Delta\theta \tag{9}$$

with $f(\theta) = 10\tan\theta$ and $\theta_0 = 30\frac{\pi}{180}$ to get

$$\Delta h = f(\theta_0 + \Delta\theta) - f(\theta_0) \approx f'(\theta_0)\Delta\theta \implies \Delta\theta \approx \frac{\Delta h}{f'(\theta_0)}$$

While this procedure is fairly reliable, it did involve an approximation. So that you could not 100% guarantee to your client's lawyer that an accuracy of 10 cm was achieved. If we use the exact formula (8), with the replacements $x \to \theta_0 + \Delta\theta$, $x_0 \to \theta_0$, $c \to \phi$,

$$f(\theta_0 + \Delta\theta) = f(\theta_0) + f'(\phi)\Delta\theta \text{ for some } \phi \text{ between } \theta_0 \text{ and } \theta_0 + \Delta\theta$$

in place of the approximate formula (2), this legality is taken care of.

$$\Delta h = f(\theta_0 + \Delta\theta) - f(\theta_0) = f'(\phi)\Delta\theta \implies \Delta\theta = \frac{\Delta h}{f'(\phi)} \text{ for some } \phi \text{ between } \theta_0 \text{ and } \theta_0 + \Delta\theta$$

Of course we do not know exactly what ϕ is. But suppose that we know that the angle was somewhere between 25° and 35°. In other words suppose that, even though we don't know precisely what our measurement error was, it was certainly no more than 5°. Then $f'(\phi) = 10\sec^2(\phi)$ must be smaller than $10\sec^2 35° < 14.91$, which means that $\frac{\Delta h}{f'(\phi)}$ must be at least $\frac{.1}{14.91}$ radians or $\frac{.1}{14.91}\frac{180}{\pi} = .38°$. A measurement error of 0.38° is certainly acceptable.

Supplement – Derivation of the Error Formulae

Define

$$E_n(x) = f(x) - f(x_0) - f'(x_0)(x - x_0) - \cdots - \tfrac{1}{n!}f^{(n)}(x_0)(x - x_0)^n$$

This is the error introduced when one approximates $f(x)$ by $f(x_0)+f'(x_0)(x-x_0)+\cdots+\frac{1}{n!}f^{(n)}(x_0)(x-x_0)^n$. We shall now prove that

$$E_n(x) = \tfrac{1}{(n+1)!}f^{(n+1)}(c)\,(x - x_0)^{n+1} \tag{$10_{\rm n}$}$$

for some c strictly between x_0 and x. **This proof is not part of the official course**. In fact, we have already used the Mean–Value Theorem to prove that $E_0(x) = f'(c)\,(x - x_0)$, for some c strictly between x_0 and x. This was the content of (8). To deal with $n \ge 1$, we need the following small generalization of the Mean–Value Theorem.

Theorem (Generalized Mean–Value Theorem) *Let the functions $F(x)$ and $G(x)$ both be defined and continuous on $a \le x \le b$ and both be differentiable on $a < x < b$. Furthermore, suppose that $G'(x) \ne 0$ for all $a < x < b$. Then, there is a number c obeying $a < c < b$ such that*

$$\frac{F(b)-F(a)}{G(b)-G(a)} = \frac{F'(c)}{G'(c)}$$

Proof: Define

$$h(x) = \big[F(b) - F(a)\big]\big[G(x) - G(a)\big] - \big[F(x) - F(a)\big]\big[G(b) - G(a)\big]$$

Observe that $h(a) = h(b) = 0$. So, by the Mean–Value Theorem, there is a number c obeying $a < c < b$ such that

$$0 = h'(c) = \big[F(b) - F(a)\big]G'(c) - F'(c)\big[G(b) - G(a)\big]$$

As $G(a) \neq G(b)$ (otherwise the Mean–Value Theorem would imply the existence of an $a < x < b$ obeying $G'(x) = 0$), we may divide by $G'(c)\big[G(b) - G(a)\big]$ which gives the desired result. ■

Proof of (10_n): To prove (10_1), that is (10_n) for $n = 1$, simply apply the Generalized Mean–Value Theorem with $F(x) = E_1(x) = f(x) - f(x_0) - f'(x_0)(x - x_0)$, $G(x) = (x - x_0)^2$, $a = x_0$ and $b = x$. Then $F(a) = G(a) = 0$, so that

$$\tfrac{F(b)}{G(b)} = \tfrac{F'(\tilde c)}{G'(\tilde c)} \implies \tfrac{f(x)-f(x_0)-f'(x_0)(x-x_0)}{(x-x_0)^2} = \tfrac{f'(\tilde c)-f'(x_0)}{2(\tilde c - x_0)}$$

for some $\tilde c$ strictly between x_0 and x. By the Mean–Value Theorem (the standard one, but with $f(x)$ replaced by $f'(x)$), $\frac{f'(\tilde c)-f'(x_0)}{\tilde c - x_0} = f''(c)$, for some c strictly between x_0 and $\tilde c$ (which forces c to also be strictly between x_0 and x). Hence

$$\tfrac{f(x)-f(x_0)-f'(x_0)(x-x_0)}{(x-x_0)^2} = \tfrac{1}{2}f''(c)$$

which is exactly (10_1).

At this stage, we know that (10_n) applies to all (sufficiently differentiable) functions for $n = 0$ and $n = 1$. To prove it for general n, we proceed by induction. That is, we assume that we already know that (10_n) applies to $n = k - 1$ for some k (as is the case for $k = 1, 2$) and that we wish to prove that it also applies to $n = k$. We apply the Generalized Mean–Value Theorem with $F(x) = E_k(x)$, $G(x) = (x - x_0)^{k+1}$, $a = x_0$ and $b = x$. Then $F(a) = G(a) = 0$, so that

$$\frac{F(b)}{G(b)} = \frac{F'(\tilde c)}{G'(\tilde c)} \implies \frac{E_k(x)}{(x - x_0)^{k+1}} = \frac{E_k'(\tilde c)}{(k+1)(\tilde c - x_0)^k} \tag{11}$$

But

$$\begin{aligned} E_k'(\tilde c) &= \tfrac{d}{dx}\Big[f(x) - f(x_0) - f'(x_0)\,(x - x_0) - \cdots - \tfrac{1}{k!}f^{(k)}(x_0)\,(x - x_0)^k\Big]_{x=\tilde c} \\ &= \Big[f'(x) - f'(x_0) - \cdots - \tfrac{1}{(k-1)!}f^{(k)}(x_0)(x - x_0)^{k-1}\Big]_{x=\tilde c} \\ &= f'(\tilde c) - f'(x_0) - \cdots - \tfrac{1}{(k-1)!}f^{(k)}(x_0)(\tilde c - x_0)^{k-1} \end{aligned} \tag{12}$$

The last expression is exactly the definition of $E_{k-1}(\tilde c)$, but for the function $f'(x)$, instead of the function $f(x)$. But we already know that (10_{k-1}) is true. So, substituting $n \to k - 1$, $f \to f'$ and $x \to \tilde c$ into (10_n), we already know that (12), i.e. $E_k'(\tilde c)$, equals

$$\tfrac{1}{(k-1+1)!}\big(f'\big)^{(k-1+1)}(c)(\tilde c - x_0)^{k-1+1} = \tfrac{1}{k!}f^{(k+1)}(c)\,(\tilde c - x_0)^k$$

for some c strictly between x_0 and $\tilde c$. Subbing this into (11) gives

$$\frac{E_k(x)}{(x - x_0)^{k+1}} = \frac{E_k'(\tilde c)}{(k+1)(\tilde c - x_0)^k} = \frac{f^{(k+1)}(c)\,(\tilde c - x_0)^k}{(k+1)\,k!\,(\tilde c - x_0)^k} = \frac{1}{(k+1)!}f^{(k+1)}(c)$$

which is exactly (10_k).

So we now know that

- if, for some k, (10_{k-1}) is true for all k times differentiable functions,
- then (10_k) is true for all $k + 1$ times differentiable functions.

Repeatedly applying this for $k = 2, 3, 4, \cdots$ (and recalling that (10_1) is true) gives (10_k) for all k. ■