Marginal Functions and Approximation

1. *The approximation formula*

If $y = f(x)$ is a differentiable function then its derivative, $y' = f'(x)$, gives the *rate of change* of the variable y with respect to the variable x. The term 'rate of change' comes from the definition of the derivative

$$f'(x_0) = \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}, \tag{1.1}$$

where $\Delta x = x - x_0$, is the change in the value of the variable x, and $\Delta y = f(x_0 + \Delta x) - f(x_0)$, is the corresponding change in the value of the variable y. Thus, the derivative is the limit, (as Δx goes to 0), of the *ratio of the changes* $\Delta y / \Delta x$. So rate-of-change comes from ratio-of-changes.

The notion of rate-of-change can be made more concrete by remembering the definition of the *limit*. Specifically, it follows from the definition of the limit that if Δx is sufficiently close to 0, then the ratio on the right-hand side of equation (1.1) is approximately equal to the value of the derivative on the left-hand side. I.e., if Δx is small, then

$$\frac{\Delta y}{\Delta x} \approx f'(x_0). \tag{1.2}$$

Now, if we multiply both sides of the approximate equality above by Δx we obtain a simple, but very important formula, that I call *the approximation formula*.

Fact 1. *If $y = f(x)$ is differentiable at $x = x_0$, and if Δx is sufficiently small, then*

$$\Delta y \approx f'(x_0) \cdot \Delta x. \tag{1.3}$$

Comments:

a. If $|\Delta x| < 1$, then the approximation in (1.3) is more accurate than the one in (1.2). (Can you say why?)
b. The quality of the estimate given by the approximation formula depends very strongly on the specific function $f(x)$, the point x_0, and the size of Δx. I will illustrate this dependence in the examples of Section 2, below.

The approximation formula may also be understood *geometrically*. Recall that the derivative $f'(x_0)$ may also be interpreted as the *slope of the tangent line* to the graph $y = f(x)$ at the point $(x_0, f(x_0))$. Using the *point-slope* formula, we find that the equation of this tangent line is

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

The *vertical distance* between the tangent line and the graph $y = f(x)$ at a point, $x_0 + \Delta x$, is the absolute value of the difference between the y-values at this point. In other words,

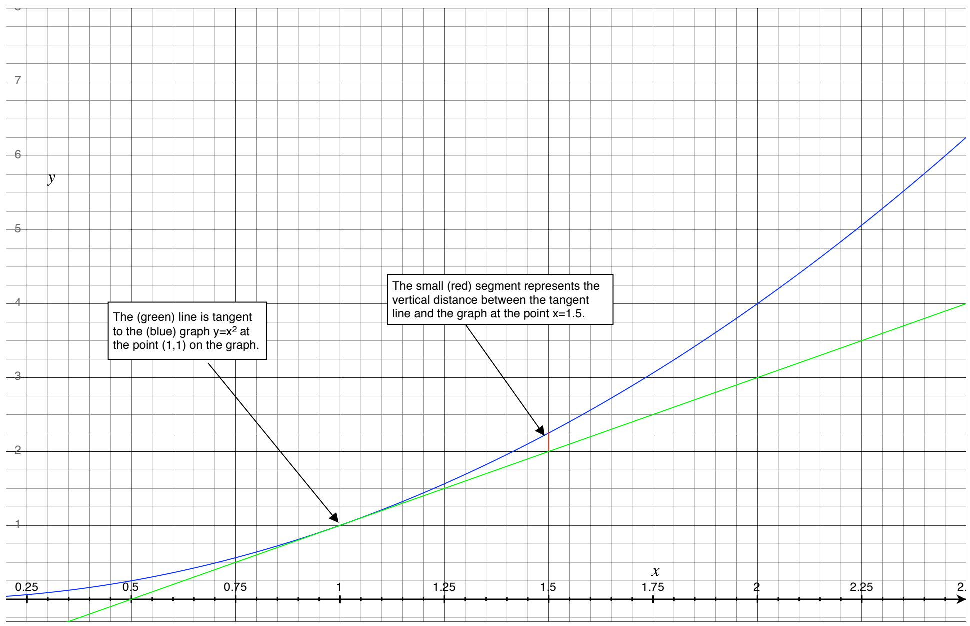

FIGURE 1. Geometric interpretation of the approximation formula.

the vertical distance between the graph and the tangent line at $x_0 + \Delta x$ is

$$\begin{aligned} \left| f(x_0+\Delta x) - \left[\, f(x_0) + f'(x_0)(x_0+\Delta x - x_0)\right] \right| &= \left| \left[\, (f(x_0+\Delta x) - f(x_0)\right] - f'(x_0)\cdot \Delta x \right| \\ &= \left| \, \Delta y - f'(x_0)\cdot \Delta x \right|. \end{aligned}$$

So, the approximation formula says that if $\Delta x = x - x_0$ is sufficiently small, then the vertical distance between the graph of the function and the tangent line (at the point $x_0 + \Delta x$) is also small. See Figure 1.

2. *Examples of the approximation formula in action*

This section contains simple examples that illustrate how the quality of the approximation in the approximation formula depends on the function $f(x)$, the starting point x_0 and the change in the x-variable, Δx. You can skip this subsection if you want, but I recommend that you don't. In fact I recommend that you redo all the computations that I do below for good measure.

Example 1. Suppose that $f(x) = \sqrt{x} = x^{1/2}$, so $f'(x) = \frac{1}{2}x^{-1/2}$. I'll apply the approximation formula to this function for two values of x_0 and several different values of Δx. First, if $x_0 = 4$, then the approximation formula reads

$$\Delta y = \sqrt{4+\Delta x} - \sqrt{x} \approx \frac{1}{2}\cdot 4^{-1/2}\cdot \Delta x = \frac{\Delta x}{4}.$$

Now, I'll compare the estimated change in the y-value provided by the approximation formula above, to the actual[†] change in the y-value for several values of Δx. For neatness'

[†]Note that this 'actual' difference is also an approximation, albeit a much more accurate one.

sake, I've collected the results in the table below, in which the left-hand column contains the different values of Δx, the middle column gives the corresponding *estimates* for Δy provided by the approximation formula and the right-hand column gives the *actual* values (rounded to 5 decimal places) of Δy.

Δx	$f'(x_0) \cdot \Delta x$	Δy
4	1	0.82843
2	0.5	0.44949
1	0.25	0.23607
0.5	0.125	0.12132
0.1	0.025	0.02485

Table 1. *Approximation formula in action: (i)* $f(x) = \sqrt{x}$ *and* $x_0 = 4$.

A couple of things may be observed. First, the estimates become more accurate as Δx gets smaller. When $\Delta x = 4$, the estimate is off by more than 0.1, and when $\Delta x = 0.1$, the estimate is off by less than 0.0002. Second, all the estimates are *too big* ‡ in this example.

Let's see what happens when we use the same function, and the same values of Δx, but with a different *starting point*, namely $x_0 = 25$. In this case the approximation formula gives

$$\Delta y = \sqrt{25 + \Delta x} - \sqrt{25} \approx \frac{1}{2} \cdot 25^{-1/2} \cdot \Delta x = \frac{\Delta x}{10},$$

and repeating the computations above produces the table

Δx	$f'(x_0) \cdot \Delta x$	Δy
4	0.4	0.38516
2	0.2	0.19615
1	0.1	0.09902
0.5	0.05	0.04975
0.1	0.01	0.00999

Table 2. *Approximation formula in action: (ii)* $f(x) = \sqrt{x}$ *and* $x_0 = 25$.

What do we see here? First of all, the same patterns we observed above still hold, namely better estimates for smaller values of Δx, and all the estimates are bigger than the actual differences. But we can also compare the results in the first table to the results in the second table, and we see that for the same values of Δx, the estimates in the second table ($x_0 = 25$) are much better§ than the corresponding estimates in the first table ($x_0 = 4$).

This example illustrated two of the dependencies that I mentioned above, namely the quality of the estimate provided by the approximation formula depends on x_0 and on Δx. The next example will show that there is also a strong dependence on the the *function*, $f(x)$.

Example 2. In this example, I'll apply the approximation formula to the function $f(x) = x^3$, and I'll generate the same table that I did in Example 1 for $x_0 = 4$, with the same values of Δx that I used before. I'll leave it to you, as an exercise, to produce the same table for $x_0 = 25$.

‡Consider the geometric interpretation of the approximation formula, to understand why.

§The estimates in the $x_0 = 25$ table are about 10 times closer to the actual values than the corresponding estimates in the $x_0 = 4$ table.

In this case, $f'(x) = 3x^2$, and for $x_0 = 4$, the approximation formula gives

$$\Delta y = (4 + \Delta x)^3 - 4^3 \approx f'(x_0)\Delta x = 3 \cdot 4^2 \cdot \Delta x = 48 \cdot \Delta x,$$

and we obtain the table

Δx	$f'(x_0) \cdot \Delta x$	Δy
4	192	448
2	96	152
1	48	61
0.5	24	27.125
0.1	4.8	4.921

Approximation formula in action: (iii) $f(x) = x^3$ *and* $x_0 = 4$.

What do we observe here? The first thing to notice that the approximations are not nearly as good in this case as they were in the first example, ($f(x) = \sqrt{x}$, $x_0 = 4$). Not until $\Delta x = 0.1$ is the distance between the estimate, $48 \cdot \Delta x$, and the actual value, $(4+\Delta x)^3 - 4^3$, less than 1. This illustrates the third dependency that I mentioned before. Namely, *the quality of the approximation depends on the* **function** *too.* The approximation formula will yield very good approximations for this function too, but we need smaller values of Δx to get them. For example, if $\Delta x = 0.01$ then

$$\Delta y = (4.01^3 - 4^3) = 0.481201 \quad \text{and} \quad 48 \cdot (0.01) = 0.48,$$

and the difference between the estimate and the actual value of Δy is about 0.0012.

There are other differences between this example and the first two above. For one thing, the estimates in this case are all *too small*, (see the footnote on the previous page). And, finally, as you should check by repeating the second half of Example 1 for yourself, if we increase x_0, then the estimated values of Δy will be worse for the same choices of Δx. (In Example 1, the estimates improved when x_0 increased.)

3. *Marginal functions and approximation in economics*

Economic activity is often described in terms of the *change* in the values of the economic functions being considered. Economists use the word '*marginal*' to describe this change. For example, the marginal revenue of a firm is defined to be the change in revenue generated by an increase in output of one unit. I.e., if $r = f(q)$ is the revenue function, where q is the firm's output and r is its revenue, then the marginal revenue is

$$mr = f(q + 1) - f(q). \tag{3.1}$$

If the function in question is differentiable, then the approximation approximation formula can be used to estimate marginal behavior. For example, applying the approximation formula, (1.3), to Equation (3.1), above, we find that

$$mr = f(q + 1) - f(q) \approx f'(q) \cdot \Delta q = f'(q) \cdot 1 = \frac{dr}{dq}. \tag{3.2}$$

In other words, the marginal revenue is approximately equal¶ to the derivative of the revenue function. All of this leads to the following definition.

¶ Strictly speaking, this statement is only accurate if the revenue function is '*well behaved*', since in general, the linear approximation formula is only accurate when Δq is 'sufficiently small', and for many functions, $\Delta q = 1$ is **not** sufficiently small.

If the revenue function $r = f(q)$ is differentiable then the **marginal revenue function** is **defined** to be the **derivative** of the revenue function, dr/dq.

This definition is only one example. In general, in the context of differential calculus applied to economics, then the word *marginal* connotes *derivative*. Some of the most common examples are listed below.

- The **marginal cost** function is the derivative of the cost function, with respect to output.
- The **marginal propensity to consume** is the derivative of the consumption function with respect to income.
- The **marginal propensity to save** is the derivative of the savings function with respect to income.
- The **marginal revenue product** is the derivative of revenue with respect to labor input (number of employees).
- The **marginal product** of capital (or labor, or x, etc.) is the derivative of output with respect to capital (or labor, or x, etc.).

Example 3. Suppose that a firm's production function is given by

$$q = 100(m+4)^{2/3},$$

where q is the firm's output, and m is the number of the firm's employees. The firm's *marginal product of labor*[‖] is

$$\frac{dq}{dm} = \frac{200}{3}(m+4)^{-1/3}.$$

Suppose that the firm's *current* workforce is $m_0 = 60$. By how much can the firm expect output to increase if they hire one more employee? According to the approximation formula

$$\Delta q \approx \left(\left. \frac{dq}{dm} \right|_{m=60} \right) \cdot \Delta m = \frac{200}{3} \cdot 64^{-1/3} \cdot 1 = \frac{50}{3} \approx 16.667.$$

In other words, if the number of employees increases from 60 to 61, then the output will approximately increase by the value of the *marginal product function* when $m = 60$. How good is this approximation? Well,

$$q(61) - q(60) = 100 \cdot 65^{2/3} - 100 \cdot 64^{2/3} = 16.62356,$$

rounded to 5 decimal places, so the approximation in this case is reasonably good — the difference between the estimate and the actual change in output is less than 0.05.

Example 4. The consumption function for a small, developing country is estimated to be

$$C = \frac{9Y^2 + 4Y + 2}{10Y + 3},$$

where C is *per-capita* consumption and Y is *per-capita* income. Both C and Y are measured in 1000's of dollars. The current per-capita income is $Y_0 = 1.2$. What will the approximate change in consumption be if per-capita income increases by $250?

The marginal propensity to consume[**] in this case is

$$\frac{dC}{dY} = \frac{(18Y+4)(10Y+3) - 10(9Y^2+4Y+2)}{(10Y+3)^2} = \frac{90Y^2 + 54Y - 8}{100Y^2 + 60Y + 9}.$$

[‖]The derivative in this example is computed using the *chain rule*.

[**]I used the quotient rule here.

The (projected) change in income is $\Delta Y = 0.25 = \frac{250}{1000}$, since Y is measured in 1000's of dollars, so the (projected) change in per-capita consumption is

$$\Delta C \approx \left(\left. \frac{dC}{dY} \right|_{Y=1.2} \right) \cdot \Delta Y = 0.828\overline{444} \cdot 0.25 = 0.207\overline{111},$$

according to the approximation formula. In dollar terms, the projected per-capita increase in consumption is about \$207.

There is a natural question to ask here. Namely, why use the approximation formula to estimate the change in the firm's output? Why not compute the change directly? In the laptop era, why compute an approximate value when the *precise* value is a few key-strokes away?

The answer has a theoretical component and a practical, 'real-world' component. The theoretical part of the answer is this. For functions like $f(x) = x^{2/3}$, $s = e^t$ or $U = \ln v$, to name a few simple examples, the values that your calculator (or laptop) produces are usually approximations themselves, albeit very good ones. In other words, there is nothing wrong with using an approximation — much of (applied) mathematics involves finding good approximations. The approximation formula is one of the most basic (and most important) approximation tools in your mathematical toolkit.

From a practical point of view, the functions that economists use to model economic reality are all produced using sophisticated statistical and mathematical tools, **from actual data**. But this data has gaps, and the approximation formula and its more sophisticated relatives are used to fill these gaps, and in certain cases *predict* future values.