Euler Methods

The interpretation of a differential equation as a formula for the rate of change of a function leads to the development of the direction field, which is a tool for visualizing solution curves without finding solution formulas. If we want to approximate solutions, rather than merely visualize them, we need a computational technique. This simplest way to do this is with *Euler's method* (named after the 18th century Swiss mathematician Leonhard Euler¹), which is based on the direction field. We will see that Euler's method has some difficulties, but we'll develop the *improved Euler method*, which is suitable for most problems.

MODEL PROBLEM 1

Approximate the solution of

$$\frac{dy}{dt} = \frac{8e^{-t}}{3+y}, \quad y(0) = 0$$

without using the solution formula.

Figure 1 shows the direction field for the differential equation of Model Problem 1. In the absence of a solution formula, some method is needed to add solution curves to the direction field plot. This is a numerical problem; in other words, some computational approximation is needed to use the data about the direction field to construct approximate solution curves.



Figure 1: The direction field for $dy/dt = (8e^{-t})/(3+y)$

1. A First Approximation

To begin, suppose we just want to approximate the solution on the interval $0 \le t \le 0.2$. There are only two pieces of information about the solution curve that we know for sure. First, it passes through the initial point (0,0). Second, its slope at the point (0,0) can be calculated using the

¹The correct pronunciation is "Oy-ler."

differential equation: $y'|_{(0,0)} = 8e^{-t}/(3+y)|_{(0,0)} = 8/3 \approx 2.667$. Graphically, this means that the solution curve has slope 8/3 at the point (0,0), as depicted by the minitangent in the direction field. The simplest way to approximate the solution on $0 \le t \le 0.2$ is to draw the straight line segment through the point (0,0) in the direction of the minitangent and ending at a point where t = 0.2. Figure 2 shows this line segment along with the direction field.



Figure 2: A line segment approximation for $dy/dt = (8e^{-t})/(3+y)$, y(0) = 0 on the interval [0,0.2], along with the direction field

The y coordinate of the right end of the line segment can easily be determined from

$$\Delta y = (\text{slope})\Delta t = (8/3)(0.2) \approx 0.533$$

and

$$y(0.2) \approx 0 + 0.533 = 0.533$$

This calculation is illustrated in Figure 3.



Figure 3: The line segment from (0,0) to t = 0.2 in the approximation for Model Problem 1

Now suppose we want to approximate the solution on the interval $0 \le t \le 0.4$. We can use the line segment just calculated up to t = 0.2, and then we can use the same procedure on the interval $0.2 \le t \le 0.4$. We sketch a line segment from the point (0.2, 0.533) to t = 0.4, using the slope of the minitangent at (0.2, 0.533), which is approximately 1.854. To find the y coordinate of the right end, we have

$$y = 0.533 + \Delta y = 0.533 + (1.854)(0.2) \approx 0.904$$

Figure 4 shows this approximation along with the direction field.

We can continue to use this same procedure to get more line segments approximating portions of the solution curve, in general using information at point n to estimate the solution at point n+1.



Figure 4: A line segment approximation for $dy/dt = 8e^{-t}/(3+y)$, y(0) = 0 on the interval [0,0.4], along with the direction field

This technique of approximating the solution on intervals $[t_n, t_n + \Delta t]$ by using line segments whose slopes are determined at $t = t_n$ is called **Euler's method**.

INSTANT EXERCISE 1

Determine the next point (t = 0.6) on the graph of the straight-line approximation to the problem

$$\frac{dy}{dt} = \frac{8e^{-t}}{3+y}, \quad y(0) = 0.$$

The full set of points obtained in this manner up to t = 1 is given in the Table 1. All values are rounded off to 3 decimal digits.

t	0	0.2	0.4	0.6	0.8	1.0
y	0	0.533	0.904	1.179	1.389	1.553
dy/dt	2.667	1.854	1.374	1.051	0.819	

Table 1: Approximate values for y and dy/dt for Euler's method applied to Model Problem 1 on the interval [0, 1], with $\Delta t = 0.2$

The upper curve in Figure 5 shows the approximate solution along with the direction field. Clearly the approximation is not perfect. A close look at the figure shows that each line segment is parallel to the direction field at its left end, but not at its right end. This is so because the solution curve is really *curved*; that is, the slope of the correct solution curve changes continuously, while the slope of the approximate solution curve changes only at specific points separated by a distance $\Delta t = 0.2$.

Normally, it is difficult to know the accuracy of an approximation. In this case, there is a solution formula:

$$y = \sqrt{25 - 16e^{-t} - 3}$$

This function is the lower curve in Figure 5.

INSTANT EXERCISE 2

Solve Model Problem 1, using the method of separation of variables.



Figure 5: The Euler approximation for $dy/dt = 8e^{-t}/(3+y)$, y(0) = 0, with $\Delta t = 0.2$, along with the direction field and the exact solution, $y = \sqrt{25 - 16e^{-t} - 3}$

2. Euler's Method in General

We will want to apply Euler's method to any problem of the form

$$\frac{dy}{dt} = f(t,y), \quad y(t_0) = y_0,$$
 (1)

where the initial point satisfies the requirements of the existence/uniqueness theorem. We assume that we want to approximate the problem for the range $t_0 \leq t \leq t_f$, with t_f any value of t larger than t_0 . (In practice, the method works equally well for $t_f \leq t \leq t_0$, with $t_f < t_0$.)

Assume that we want to divide the interval $[t_0, t_f]$ into N equal subdivisions, and that we want to approximate y(t) at each of the subdivision points by a quantity y_n . The step size is

$$\Delta t = \frac{t_f - t_0}{N}$$

and the subdivision points are

$$t_n = t_0 + n\Delta t, \quad n = 0, 1, \cdots, N.$$
 (2)

In particular, the last point is $t_N = t_f$. Notice that the *t* coordinates of the approximation points are all determined in advance by the choices of initial condition, final point, and number of steps. The *y* coordinates of the approximation points must be determined sequentially, from left to right, by assuming that the slope of each line segment is that of the minitangent at the left endpoint.

Initially, we know only y_0 . On the first interval $[t_0, t_1]$, we approximate the slope m_1 by

$$m_1 \approx m_{1e} = f(t_0, y_0)$$

where the extra subscript e indicates that the slope is that of Euler's method rather than some other method. A line segment on an interval of length Δt yields the approximate change in y of

$$\Delta y_1 = m_1 \Delta t.$$

The right endpoint of the line segment is the point $(t_1, y_0 + \Delta y_1)$. We want y_1 to approximate the y coordinate at the point t_1 , so we choose

$$y_1 = y_0 + \Delta y_1.$$

The same calculation works for the next subinterval and, in turn, all the others.

We can use a pair of formulas to calculate any y_{n+1} from the previously calculated y_n . We first calculate a slope,

$$m_{n+1,e} = f(t_n, y_n) \tag{3}$$

for interval n + 1 and then use it to calculate a linear approximation for the point at the right end of the interval,

$$y_{n+1} = y_n + m_{n+1,e}\Delta t.$$
 (4)

The full Euler approximation is obtained by calculating all points (t_n, y_n) with Equations (2–4), starting with n = 0, and then connecting the points with line segments. See Figure 6.



Figure 6: A line segment in Euler's method for the general case

3. Improving the Results

The results of the calculations do not seem to be very good, and a careful examination of Figure 5 shows why. The point (0,0) is on the correct solution curve, but the next approximation point, (0.2, 0.533), is not. This point is on a different solution curve. If we could avoid making any further error beyond t = 0.2, the numerical approximation would follow a solution curve from then on, but it would not be the *correct* solution curve. The amount of error ultimately caused by the error in the first approximation step can become larger or smaller at later times, depending on the relationship between the different solution curves.² The real situation is more complicated, because additional error is introduced at each step of the approximation.

As we have noted, the slope of the correct solution curve changes continuously, while the slope of the approximate solution curve changes only at specific points separated by a distance Δt . It seems reasonable to guess that we can get a better approximation by using a smaller value of Δt . This means that the slope of the approximate solution curve will be corrected more often. Instead of using subintervals of width 0.2, we could try using subintervals of width 0.1 or less. We cannot just make Δt extremely small without introducing other difficulties. Using smaller subintervals means that the slope is changed more often, but at a cost of more calculations. (Taking $\Delta t = 10^{-9}$, for example, would require calculation of 1 billion data points to obtain the graph.) The approximations using the values 0.2 and 0.1 are illustrated in Figure 7, along with the correct solution. The circles near the correct solution curve were obtained with $\Delta t = 0.025$; these are almost accurate to visual standards.

²You can find more complete information about error propagation in many differential equations books.



Figure 7: The Euler approximations using $\Delta t = 0.2$ and $\Delta t = 0.1$, for $dy/dt = 8e^{-t}/(3+y)$, y(0) = 0, along with the correct solution $y = \sqrt{25 - 16e^{-t}} - 3$ and data points obtained with $\Delta t = 0.025$

The error decreases to roughly half when the step size is cut in half. (See Table 2.) This is a general property of Euler's method.

	Approximation				Solution	Error			
t	Δt				y(t)	Δt			
	0.2	0.1	0.05	0.025		0.2	0.1	0.05	0.025
0.2	0.533	0.488	0.468	0.459	0.450	0.083	0.038	0.018	0.009
0.4	0.904	0.837	0.807	0.792	0.778	0.126	0.059	0.029	0.014
0.6	1.179	1.099	1.062	1.045	1.027	0.152	0.072	0.035	0.018
0.8	1.389	1.301	1.259	1.240	1.220	0.169	0.081	0.039	0.020
1.0	1.553	1.458	1.414	1.393	1.372	0.181	0.086	0.042	0.021

Table 2: Comparison of results for Euler's method applied to Model Problem 1 on the interval [0, 1] with several different step sizes

4. A Test Problem

Consider the family of problems

$$y\frac{dy}{dt} = \frac{\pi}{2}\cos\pi t, \quad y(0) = y_0 \approx 1.$$
 (5)

This problem has solution

$$y = \sqrt{y_0^2 + \sin \pi t}.\tag{6}$$

INSTANT EXERCISE 3

Derive the solution formula for the test problem (5).

The test problem is interesting for two reasons. First, the solution is periodic, with $y(n) = y_0$ for any positive integer n. This gives us a convenient check on quantitative accuracy over a long interval. Second, the interval of existence of the solution depends critically on y_0 . If $y_0 > 1$, then the solution exists for all t. If $y_0 < 1$, the solution ceases to exist a bit sooner than t = 3/2. Error

in the method may lead to failure in these qualitative results. It could be that solutions will cease to exist for some y_0 values greater than 1 or will continue to exist for some y_0 values less than 1. We cannot expect a numerical method to give an exact answer, but we must be alert to the danger that a numerical answer is qualitatively wrong.

Figure 8 shows the exact solution to the test problem with $y_0 = 1$ on the interval [0,2] along with the Euler approximations using 100 steps and 200 steps. The exact solution drops to exactly 0 at t = 1.5 and then rebounds. The numerical approximations are visually acceptable up to about t = 0.5, but soon afterward they drift away from the exact solution. At t = 1.5, the error is qualitative as well as quantitative. Curiously, the approximate solutions come back close to the exact solution after that, although the same problem will occur at t = 3.5, the next time the exact solution reaches 0.



Figure 8: The Euler approximations for $y dy/dt = (\pi/2) \cos \pi t$, y(0) = 1, using 100 and 200 time steps on the interval [0, 2], along with the exact solution (dashed)

Now suppose $y_0 = 0.9$. In this case, the solution should cease to exist when y reaches 0 at $t \approx 1.30$. Figure 9 shows the approximations on the interval [0, 2] with 200 and 400 time steps. The solution does not stop at t = 1.30, but it clearly exhibits some strange behavior near that point. The approximate curves quickly move away from y = 0, but the direction in which they move depends on the step size. What is happening?

We can draw one conclusion simply from the observation that doubling the number of steps makes a drastic change in the numerical result. Doubling the number of steps should make the approximation look more like the correct solution, in which case we should not be seeing a qualitative difference between the two approximations. This should make us suspicious even if we do not have a solution formula or any other knowledge. This sort of behavior indicates that the method is not working, and it suggests that there is something unusual about the solution at that point.

Careful examination of the differential equation also helps explain the problem. In the form dy/dt = f(t, y), the differential equation is

$$\frac{dy}{dt} = \frac{\pi \cos \pi t}{2y}$$

The function that gives the slope is unbounded as $y \to 0$. As y becomes small, the calculated slopes become large. If y is negative at such a point, we will get a large negative slope; otherwise we will get a large positive slope. A large slope magnifies the error made in using a line segment to approximate the solution curve. Each of the sudden jumps in Figure 9 is the result of one slope calculation that yields a result large in magnitude. It is simply a coincidence that one of these calculations was done at a point with positive y and the other with negative y. The errors in these calculations are so large that the results beyond these points are meaningless.



Figure 9: The Euler approximations for $y dy/dt = (\pi/2) \cos(\pi t)$, y(0) = 0.9 using 200 and 400 time steps on the interval [0, 2]

The importance of this example can hardly be overstated. Computers are a wonderful tool for approximating and visualizing the solutions of differential equations. However, computers cannot tell whether something is wrong with their calculations unless they are programmed to do so. Professional numerical differential equation solvers generally avoid difficulties such as these, but no numerical scheme is so good as to be immune to trouble. The educated user needs to be able to look at numerical results and spot indications of trouble. Difficulties such as we have seen here are caused by a combination of method error³ and *ill-conditioning* of the differential equation.⁴

5. The Improved Euler Method

Euler's method is the most conceptually clear way to approximate the solutions of differential equations, but we've seen that it doesn't always work well and we know why: it fails to account for the change in the slope of solution curves on a time interval. The key to improving on Euler's method is to keep the update formula (4), but improve on the slope approximation formula (3). We will use

$$y_{n+1} = y_n + m_{n+1,i} \Delta t, \tag{7}$$

where $m_{n+1,i}$ is an "improved Euler" slope used for interval n+1.

The question, of course, is how do we get a better approximation of the slope when we have so little information to work with? If we could calculate the slope at any point on the solution curve, then the best thing to do would be to average the slopes at the points (t_n, y_n) and (t_{n+1}, y_{n+1}) and use that average as the slope for segment n, which connects these two points. We can't do that because we can't calculate y_{n+1} until *after* we have chosen the approximate slope for the segment. The improved Euler method is based on the simple idea that we can use Euler's method to approximate y_{n+1} and then use that approximate value to approximate the slope at the right end of the line segment. The formulas needed to calculate y_{n+1} are then

$$m_e = f(t_n, y_n),\tag{8}$$

$$y^+ = y_n + m_e \Delta t, \tag{9}$$

³In the context of numerical approximations, the word *error* does not mean that the approximation scheme or computer is making a mistake. Rather, it refers to errors that are a natural consequence of the numerical approximation. An example is the rounbd-off error that occurs when a computer uses a finite number of digits for π .

⁴These topics are explored further in books on numerical methods for ordinary differential equations.

$$m^+ = f(t_{n+1}, y^+),$$
 (10)

$$m = 0.5(m_e + m^+), \tag{11}$$

and

$$y_{n+1} = y_n + m\,\Delta t,\tag{12}$$

where we have suppressed the unnecessary subscripts for the improved Euler slope and used the plus sign to indicate approximations for interval n + 1 based on the Euler approximation for y_{n+1} .

To calculate the first point in Model Problem 1, we have n = 0, $t_n = 0$, and $y_0 = 0$. Taking $\Delta t = 0.2$ as we did with Euler's method, we have⁵

$$m_e = f(t_n, y_n) = \frac{8e^0}{3+0} = \frac{8}{3} \approx 2.6667,$$

$$y^+ = y_n + m_e \Delta t = 0 + \left(\frac{8}{3}\right)(0.2) \approx 0.5333,$$

$$m^+ = f(t_{n+1}, y^+) = \frac{8e^{-0.2}}{3+0.5333} = 1.8537,$$

$$m = 0.5(m_e + m^+) = 0.5(2.6667 + 1.8537) = 2.2602.$$

and

$$y_1 = y_0 + m\,\Delta t = 0.4520.$$

Compared to the correct answer of 0.4500, this is an error of only 0.0020. From Table 2 and the rule that doubling the Euler step size halves the error, we can estimate that Euler's method needs $\Delta t = 0.005$ to achieve a similar accuracy. What requires 40 steps of Euler's method to accomplish can be done with just one step of the improved Euler method.

INSTANT EXERCISE 4

Use the improved Euler method to calculate $y_2 \approx y(0.4)$ from the starting point (0.2, 0.4520).

Is the improvement in the improved Euler method always this good? No. Sometimes it is *much* better than 40 to 1! Whether the improvement is small or large depends on the differential equation. Regardless of the problem, the improved Euler method has the general property that cutting the step size in half cuts the error by a factor of 4 rather than a factor of 2, as with Euler's method.

Example 1

Obtain a numerical approximation for the solution of

$$y\frac{dy}{dt} = \frac{\pi}{2}\cos\pi t, \quad y(0) = 1.01$$
 (13)

with error less than 0.001 at each point.

First note that the performance of Euler's method on this problem is not impressive. Although the solution,

$$y = \sqrt{1.0201 + \sin \pi t}$$

is defined for all t, Euler's method has difficulty approximating this solution when it is near its minimum point at (1.5, 0.1418). Figure 10 shows the Euler approximations with 100 and 200 time steps on the interval [0, 2].

⁵Most books combine formulas (8)–(12) into one single ugly formula. This is fine if you want fewer formulas and don't mind more computational errors. Whether working problems by hand or coding them on a computer, you will make far fewer errors with a sequence of simple formulas than with one complicated formula.



Figure 10: The Euler approximations for Model Problem 1 using 100 and 200 steps (solid) along with the exact solution (dashed)

By trial and error, it turns out that Euler's method requires N = 2047 over the interval [0, 1.5] to achieve an error less than 0.01. Since doubling the number of steps roughly halves the error, we can expect to have to increase the number of steps by a factor of 10 to reduce the error by a factor of 10 so as to get the error down to 0.001. By this estimate, we would need to use more than 20,000 steps to achieve the desired accuracy with Euler's method!

Figure 11 shows the improved Euler approximation for (13) with just 20 time steps as points, along with the exact solution.



Figure 11: The improved Euler approximation with 20 steps (points) for Model Problem 2.6 along with the exact solution

We noted earlier that Euler's method achieves an error less than 0.01 at t = 1.5 only with 2047 steps or more. In contrast, the improved Euler method achieves the same accuracy with only 33 steps. Since each step requires two function evaluations, the computer time required for this level of accuracy with the improved Euler method is about 3.2% of that required for Euler's method. The comparison between the methods is even more striking when one looks at the effort required to achieve a high degree of accuracy. Whereas Euler's method was estimated to require more than 20,000 steps to achieve an error less than 0.001, trial and error shows that the improved Euler method accomplishes this goal with just 91 steps, or 182 function evaluations, which is less than 1% of the requirement for Euler's method. Methods that use the update formula $y_{n+1} = y_n + m_{n+1} \Delta t$ are called $Runge^6$ -Kutta methods. The improved Euler method is excellent, but it is far from the best. Most people who write their own simple computer codes use the "classical" Runge-Kutta method, known generally as RK4, which averages four different slopes to calculate m. Many professional ODE solvers use a variant of RK4, called RKF4, which uses additional calculations to monitor the error and modify the step size if the error is too large. For reference, RK4 is defined by

$$m_1 = f(t_n, y_n),\tag{14}$$

$$m_2 = f(t_n + 0.5\,\Delta t, y_n + 0.5\,m_1\,\Delta t),\tag{15}$$

$$m_3 = f(t_n + 0.5\,\Delta t, y_n + 0.5\,m_2\,\Delta t),\tag{16}$$

$$m_4 = f(t_{n+1}, y_n + m_3 \,\Delta t),\tag{17}$$

and

$$m = \frac{m_1 + 2m_2 + 2m_3 + m_4}{6}.$$
(18)

The similarity to the improved Euler method is not too hard to see: m_1 and m_4 are approximate slopes at the beginning and end of the interval and m_2 and m_3 are two different approximations for the slope at the midpoint. The weights 1/6, 1/3, 1/3, and 1/6 give just the right average for a very high degree of accuracy. While cutting the step size decreases the error by a factor of 2 for Euler's method and 4 for the improved Euler method, it decreases the error by a factor of 16 for the RK4 method.

EXERCISES

Note: Exercises with an asterisk in front of the number require a computer.

1. Consider the problem

$$\frac{dy}{dt} + y = 2t + 2, \quad y(0) = 2,$$

which has solution $y = 2t + 2e^{-t}$.

a. Use Euler's method (do not use a computer program) with step size 0.1 to approximate y(0.1), y(0.2), y(0.3), and y(0.4).

b. Use the improved Euler method (do not use a computer program) with step size 0.2 to approximate y(0.2) and y(0.4) for the problem.

c. Compare the results with the exact solution.

2. Consider the initial-value problem

$$x\frac{dy}{dx} = 2y, \quad y(1) = 2.$$

a. Find the exact solution.

b. Use Euler's method (do not use a computer program) with step size 0.25 to approximate y(1.25), y(1.5), y(1.75), and y(2).

c. Use the improved Euler method (do not use a computer program) with step size 0.5 to approximate y(1.5) and y(2) for the problem.

d. Compare the results with the exact solution.

⁶Run-guh

***3.** Consider the initial-value problem

$$\frac{dy}{dt} = 2ty^2, \quad y(0) = 5.$$

a. Find the exact solution.

b. Use Euler's method with step size of 0.02 to approximate y(0.4). Compare the approximation with the correct solution and determine the magnitude of the error.

c. Repeat part b, using step sizes of 0.01 and 0.005. What effect does halving the step size have on the error?

d. Use the improved Euler method with step size of 0.05 to approximate y(0.4). Compare the approximation with the correct solution and determine the magnitude of the error. Compare this result with parts b and c. e. Repeat part d using step sizes of 0.025 and 0.0125. What effect does halving the step size have on the error?

f. Plot the approximations and the exact solution.

*4. Consider the initial-value problem

$$\frac{dy}{dt} = \frac{2\sin t}{1-y}, \quad y(0) = -1$$

a. Find the exact solution.

b. Use Euler's method with 18 steps on the interval [0, 1.8].

c. Repeat part b, using 36 steps and 72 steps.

d. Plot the approximations along with the exact solution. Can you explain the graph?

e. Use the improved Euler method with 9 steps on the interval [0, 1.8].

f. Repeat part a, using 18 steps and 36 steps.

g. Plot the approximations along with the exact solution.

h. Compare with the results of parts e and f with those of parts b and c.

INSTANT EXERCISE SOLUTIONS

1. The current point is $t_2 = 0.4$, $y_2 = 0.904$. The slope at this point is

$$\frac{8e^{-t_2}}{3+y_2} \approx 1.374.$$

The change in y is determined by

$$\Delta y \approx 1.374 \Delta t = 0.275.$$

(Note that we are rounding at the third digit because previous values were already rounded at that point.) Thus, $y_3 = 1.179$.

2. The equation in separated form is

$$(3+y)\frac{dy}{dt} = 8e^{-t}.$$

Integration and substitution yield

$$\int (3+y) \, dy = 8 \int e^{-t} \, dt,$$
$$3y + \frac{y^2}{2} = -8e^{-t} + C.$$

or

This implicit solution formula can be solved explicitly for y by completing the square. First multiply both sides by 2 to get

$$y^2 + 6y = 2C - 16e^{-t}.$$

Adding 9 to both sides completes the square:

$$y^2 + 6y + 9 = (2C + 9) - 16e^{-t},$$

 or

$$(y+3)^2 = A - 16e^{-t},$$

where A is an arbitrary constant. This is a convenient point at which to determine the constant. The solution passes through the origin if $A = 3^2 + 16 = 25$. Thus,

$$(y+3)^2 = 25 - 16e^{-t}.$$

There are two possible solutions to this last equation:

$$y = -3 \pm \sqrt{25 - 16e^{-t}}.$$

The positive root is the correct one because y(0) > -3. Thus,

$$y = \sqrt{25 - 16e^{-t}} - 3.$$

3. The equation is already separated, but it is convenient to multiply by 2 before integrating. We have

$$\int 2y \, dy = \int \pi \cos \pi t \, dt.$$

The integration yields

$$y^2 = \sin \pi t + C,$$

and the initial condition yields $C = y_0^2$. Since $y_0 > 0$, the correct square root is the positive one, so the result is Equation (6).

4. $y_2 = 0.7817$