MAT 22B - Lecture Notes

12 August 2015

Euler's Method

Euler's method is a way to get a computer to estimate the solution of an initial value problem. It's a very straightforward method - if you sat down and tried to make one up in the simplest way possible, you'd probably come up with Euler's method.

First we should say more precisely what we mean when we say that the computer should "estimate the solution" of an IVP. Generally, IVP's don't have solutions that are given by tidy formulas for $y(t)$. Hence, when solving an IVP on a computer, the output is generally a list of y-values, computed at some specified list of t-values. So, in order for a computer to know what to do, we need to give it an IVP (i.e. an ODE $y' = f(t, y)$ and initial condition $y(t_0) = y_0$), and specify at which t-values we want the solution to be evaluated. For convenience, we'll denote the list of t-values $(t_1, \ldots, t_N)$

The idea behind Euler's method is a kind of "bootstrapping". To start, we know that the point (t_0, y_0) is on the solution curve. Moreover, we know the *slope* of the solution curve at this point - it's just $y' = f(t_0, y_0)$. These two pieces of information tell us line which is tangent to the solution curve at (t_0, y_0). Assuming that the step size $t_1 - t_0$ is small, we estimate that the value of the actual solution is close enough to the value of the tangent line. Putting this into formulas, we have

$$y_1 = y_0 + (t_1 - t_0)f(t_0, y_0)$$

where $y_1 \approx y(t_1)$. Then we just repeat this process: for any n between 0 and $N - 1$, we set

$$y_{n+1} = y_n + (t_{n+1} - t_n)f(t_n, y_n)$$

Given the initial condition (t_0, y_0) and the sequence $(t_1, \ldots, t_N)$, the above formula determines the sequence $(y_1, \ldots, y_N)$, which is our estimated solution.

Let's see how this looks in an example. Consider the IVP

$$y' = y, \quad y(0) = 1$$

That is, $y' = f(t, y)$ with $f(t, y) = y$. We know that the solution is $y(t) = e^t$. Let's see what Euler's method tells us. To do that, we need to specify a list of t-values. For simplicity, let's just take

$(t_1, \ldots, t_N) = \left(\frac{1}{N}, \frac{2}{N}, \ldots, \frac{N-1}{N}, 1\right)$, so that the step size $(t_{n+1} - t_n) = \frac{1}{N}$ for all n. For now, we'll leave N as an unspecified large integer.

What does Euler's method say is the value of the solution at time t_1? Well, using the formula we derived above, it should be

$$\begin{aligned} y_1 &= y_0 + (t_1 - t_0) f(t_0, y_0) \\ &= 1 + \frac{1}{N} f(0, 1) = 1 + \frac{1}{N} \end{aligned}$$

We know that the solution should be increasing, and $1 + \frac{1}{N} > 1$, so we're not going completely crazy. Now what about the n^{th} term? We can do a little simplification right off the bat:

$$\begin{aligned} y_{n+1} &= y_n + \frac{1}{N} f(t_n, y_n) \\ &= y_n + \frac{1}{N} y_n \\ &= \left(1 + \frac{1}{N}\right) y_n \end{aligned}$$

Hey! That's pretty neat: to get to the next y-value, we simply multiply by $1 + \frac{1}{N}$. With a little thought, you can convince yourself that for any n, we have $y_n = \left(1 + \frac{1}{N}\right)^n$.

Now let's get a bit of a picture of how well this estimation mimics the actual solution. According to our formula, we have that the final value is given by $y_N = \left(1 + \frac{1}{N}\right)^N$. Taking the limit as $N \to \infty$ (that is, the number of steps becomes infinite), we can recall from calculus that

$$\lim_{N \to \infty} \left(1 + \frac{1}{N}\right)^N = e$$

which exactly matches the value of the exact solution $y(1) = e^1 = e$!

The above calculation is not meant to be representative of how you use Euler's method in the real world, but rather to show you how the method produces something that reasonably recreates the actual solution to an IVP. In practice, Euler's method is not used very much. This is because being so simple, it is highly prone to errors. To see this, try modifying the above calculation to estimate $y(2)$ and $y(3)$, and see how the difference between the estimated and exact solutions grows with time. The reason this occurs is that for certain IVP's, errors compound - underestimating y_1 leads to underestimating the *slope* at (t_1, y_1) as well, and so leads to a greater underestimation of y_2, and so on.

The most commonly used method in practice is called the Runge-Kutta method, which essentially uses some tricks to get a better value to use for the slope between two points on the solution curve. Interestingly, the method was developed around 1900, way before the advent of computers. The motivation was theoretical - proving that the estimated solution converges in the limit of step size going to zero is one way to prove that an IVP has a solution.

Second Order, Linear, Homogeneous Equations with Constant Coefficients

Phew, that's a mouthful. Let's unpack what this means. A second order linear equation has the form

$$a_2(t)y'' + a_1(t)y' + a_0(t)y = g(t)$$

Saying that the equation is *homogeneous* means that the right-hand side, $g(t)$ is equal to zero. Finally, saying that the equation has *constant coefficients* means that the coefficient functions a_0, a_1, a_2 do not depend on time - they are just constants. So, the equations we are talking about have the form

$$ay'' + by' + cy = 0$$

The prototypical example of such an equation is a mass on a spring with damping. If $y(t)$ denotes the position of the mass away from equilibrium at time t, then the spring force is (by Hooke's law) equal to $-ky$, where $k > 0$ is the spring constant, and the damping force is $-b\frac{dy}{dt}$, where $b > 0$ is called the damping coefficient.

$$\begin{aligned} F &= ma \\ -ky - b\frac{dy}{dt} &= m\frac{d^2y}{dt^2} \end{aligned}$$

Rearranging, we get

$$m\frac{d^2y}{dt^2} + b\frac{dy}{dt} + ky = 0$$

So the equation of motion is second order, linear, and homogeneous, with constant coefficients. If you're not sick of masses on springs by the end of the class, I'll give you a donut. On a spring.

As we've mentioned before, the fact that these equation involves a 2nd derivative means that its *general solution* will have two free parameters. To see what that looks like, and to get a hint at the general strategy for solving these equations, let's consider a simple example:

$$y'' - 4y = 0$$

A solution to this equation has the property that its second derivative is four times itself. An example of a function that comes back to itself after differentiating twice is the exponential function - and, if we choose the right one, we can get the 4 to come out also. Consider the function $y_1 = e^{2t}$. Then $y_1' = 2e^{2t}$ and $y_1'' = 4e^{2t} = 4y$. So we've found a solution! Now, notice that $y_2 = e^{-2t}$ works as well - the first derivative is $y_2' = -2e^{-2t}$, and $y_2'' = 4e^{-2t}$. The fact that made both of these functions work is the fact that $2^2 = (-2)^2 = 4$.

So far we have found two solutions. But what is the *general solution*? That is, how do we write *all* solutions of $y'' - 4y = 0$ in one expression?

The answer relies on two key properties of the DE: *linearity* and *homogeneity*. Consider $y = y_1 + y_2$. Is this also a solution of the ODE? Let's check

$$\begin{aligned} y'' - 4y &= (y_1 + y_2)'' - 4(y_1 + y_2) \\ &= \left(y_1'' - 4y_1\right) + \left(y_2'' - 4y_2\right) \\ &= 0 + 0 = 0 \end{aligned}$$

and indeed, it is! The step from the first line to the second was made possible by the fact that the ODE is linear. The fact that it is homogeneous gives us that each term in parentheses on the second line is zero, and so they add to zero, making the sum $y_1 + y_2$ a solution also. If there were a nonzero right-hand side $g(t)$, then the last line would read $g(t) + g(t) = 2g(t) \neq g(t)$, and $y_1 + y_2$ would not be a solution.

The argument outlined above establishes that for *any* real constants c_1 and c_2, the function $y = c_1 y_1 + c_2 y_2$ is also a solution to the ODE. Physicists call this the princciple of superposition. In math speak, this means that the set of solutions to the ODE makes up a vector space. Whatever words you use, you should know that it's true and why it's true. What's more, these two functions happen to satisfy a certain property that means that *any* solution to the ODE $y'' - 4y$ has the form $c_1 y_1 + c_2 y_2$, and so we can rightly call the expression $c_1 y_1 + c_2 y_2$ the general solution to the ODE. We'll talk about the property I referred to in detail a few lectures from now.

The Characteristic Equation

The key thing that made both e^{2t} and e^{-2t} is that both 2 and -2 were solutions to the equation $r^2 - 4 = 0$. We call this equation the *characteristic equation* of the ODE $y'' - 4y = 0$. To see where it came from and how it generalizes, let's consider the general form of a second order, linear, homogeneous ODE with constant coefficients:

$$ay'' + by' + cy = 0$$

Now what we'll do, since exponential functions have been so good to us this far, is *guess* that there's a solution of the form $y = e^{rt}$ for some constant r. If we plug in this guess, we find

$$\begin{aligned} ay'' + by' + cy &= a\left(e^{rt}\right)'' + b\left(e^{rt}\right)' + c\left(e^{rt}\right) \\ &= ar^2\left(e^{rt}\right) + br\left(e^{rt}\right) + c\left(e^{rt}\right) \\ &= \left(ar^2 + br + c\right)e^{rt} \end{aligned}$$

Now, if e^{rt} is to be a solution, the above expression must be equal to zero. In other words, either of the factors $(ar^2 + br + c)$ or e^{rt} must be equal to zero for all t. However, e^{rt} is *always* nonzero, regardless of r and t. So our only hope is to pick r to be a solution of the equation

$$ar^2 + br + c = 0$$

This is called the characterisitic equation of the ODE $ay'' + by' + cy = 0$. By the quadratic formula, its solutions are

$$r_{\pm} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

and as we know from algebra, these can be either real and distinct, real and repeated, or complex. We'll deal with each of these cases separately, but with an eye towards them being secretly the same thing.

Real and Distinct Roots

This is the simplest case. When the characteristic equation has two distinct real roots r_+ and r_-, we set $y_1 = e^{r_+ t}$ and $y_2 = e^{r_- t}$, and the general solution is $c_1 y_1 + c_2 y_2$. Again, in a few lectures we'll see that in this case, the functions y_1 and y_2 satisfy a property that means that this expression captures all possible solutions.

Examples

$$y'' + y' - 2y = 0 \implies y = c_1 e^t + c_2 e^{-2t}$$
$$y'' - 5y' + 6y = 0 \implies y = c_1 e^{2t} + c_2 e^{3t}$$
$$2y'' + 3y' + y = 0 \implies y = c_1 e^{-t} + c_2 e^{-t/2}$$