NSDE 1: LECTURE 4

TYRONE REES[*]

Recap

$$u' = f(t,u), \qquad u(t_0) = u_0.$$

We defined the **global error** as

$$e_n = u(t_n) - U_n$$

and the truncation error as

$$T_n = \frac{(t_{n+1} - u(t_n))}{h} - \Phi(t_n, t_{n+1}, U_n, U_{n+1}; h).$$

For Euler's method:

$$T_n - \frac{h}{2}u''(\xi_n)$$

We saw last time that

$$|e_n| \leq \frac{T}{L}\left[e^{L(t_n - t_0)} - 1\right]$$

Finally, note that

$$T = \max_n |T_n| = \max_n \frac{h}{2}|u''(\xi_n)| \leq \frac{1}{2}hM_2$$

and so

$$|e_n| \leq \frac{M_2}{2L}(e^{L(t_n-t_0)} - 1)h \leq \text{Const.}h$$

`euler_half_step.m`

Hence, if you half the step size, you (at least) half the error.

θ-methods

A similar (but more messy) analysis can be done for $\theta-$methods. Here we get that

$$|e_n| \leq \frac{h}{L}\left\{\left|\frac{1}{2} - \theta\right| M_2 + \frac{1}{3}hM_3\right\}\left[\mathrm{e}^{\frac{L(t_n - t_0)}{1-\theta L h}} - 1\right],$$

where $M_3 = \max_{t\in[t_0,t_m]} |u'''(t)|$.

[*]Rutherford Appleton Laboratory, Chilton, Didcot, UK, tyrone.rees@stfc.ac.uk

As expected, the error of explicit Euler is recovered if we take $\theta = 0$. We also see that the special choice of $\theta = 1/2$ – the trapezium rule method – gives second order accurate method. This explains the behaviour in the numerical example.

Question: where did M_3 go for Euler's method?

Recall that when we derived the truncation error, we cut off the Taylor series at the u'' term. We could equally have written

$$T_n = \frac{h}{2}u''(t_n) + \frac{h^2}{6}u'''(\hat{\xi}_n), \qquad \hat{\xi} \in [t_n, t_{n+1}.$$

Check that, if we'd included this term, we obtain the bound for the θ-method.

One-step methods

Euler's method is an example of what's known as a one-step method. The general form of a one-step method is:

$$\begin{aligned} U_0 &= u(t_0) \\ U_{n+1} &= U_n + h\Phi(t_n, U_n; h) \end{aligned}$$

Note that all that appears in the right hand side is U_n.

Examples

Euler's method:

$$\Phi(t_n, U_n; h) = f(t_n, U_n)$$

The trapezium rule is **not** a one-step method:

$$\Phi = \frac{1}{2}\left(f(t_n, U_n) + f(t_{n+1}, U_{n+1})\right)$$

We can, however, get what should be a better method, which is **explicit**, by replacing U_{n+1} itself by it's Euler approximation, $U_n + hf(t_n, U_n)$. This is called **improved Euler's method**.

`improved_euler.m`

For any one-step method, if Φ satisfies a Lipschitz condition:

$$|\Phi(t, u; h) - \Phi(t, v; h)| \leq L|u - v|,$$

then the same analysis we did for Euler's method will go through here also. We therefore also have the concept of **Truncation error** here:

$$T_n = \frac{u(t_{n+1}) - u(t_n)}{h} - \Phi(t_n, u(t_n); h),$$

and, as before, if $T = \max_n |T_n|$, then

$$|U_n - u(t_n)| \leq (e^{L(t_n - t_0)} - 1)\frac{T}{L}$$

In order for the error to vanish with a small enough step size, we therefore require that $T_n \to 0$ as $h \to 0$ and $n \to \infty$, with $nh = t_n - t_0$.

Consistency: Definition We say that a one-step method is **consistent** if

$$\lim_{h\to 0, n\to\infty, nh=t_n-t_0} T_n = 0$$

Now, since $\Phi(\cdot,\cdot;\cdot)$ and $y'(\cdot)$ are continuous, we know that

$$\lim_{h\to 0} T_n = u'(t_n) - \Phi(t_n, u(t_n); 0).$$

Therefore a one-step method is consistent if and only if

$$\Phi(t_n, u(t_n); 0)) = f(t, u).$$

In general, we are not restricted to one extra point in the region $[t_n, t_n + h]$; we can evaluate the function at as many intermediate points as we like. We're only interested in explicit methods, so we need to also have a way of approximating the solution at those points (as $u(t_n + ah)$ will not be avaliable). This leads to a general family of methods known as Runge-Kutta methods.

Runge-Kutta R-stage method

$$\begin{aligned}
U_{n+1} &= U_n + h\Phi(t_n, U_n; h) \\
\Phi(x, y; h) &= \sum_{r=1}^{R} c_r k_r \\
k_1 &= f(t_n, U_n) \\
k_r &= f\left(t_n + a_r\, h, U_n + h\sum_{s=1}^{r-1} b_{rs} k_s, \quad r = 2, \ldots, R\right) \\
a_r &= \sum_{s=1}^{r-1} b_{rs}, \quad r = 2, \ldots, R
\end{aligned}$$

This is usually written in the form of a Butcher table:

$$\begin{array}{c|c} a = B\mathbf{1} & B \\ \hline & c^T \end{array}$$

R=1

If $R = 1$, then we get back our old friend Euler's method.

R = 2

For $R = 2$, we have more choice. Now we want to find a method such that

$$U_{n+1} = U_n + h(c_1 k_i + c_2 k_2),$$

where

$$k_1 = f(t_n, U_n)$$
$$k_2 = f(t_n + a_2 h, U_n + b_{21} h k_1)$$

What values of c_1, c_2, a_2, and b_{21} make sense?

To be consistent, we need that $\Phi(t_n, U_n; h) = f(t_n, U_n)$, i.e.

$$c_1 f(t_n, U_n) + c_2 f(t_n, U_n) = f(t_n, U_n) \iff c_1 + c_2 = 1$$

So, given c_2, we can obtain c_1, but how should we choose c_2, b_{21} and a_2. We try to make the order of the method as high as possible.

Recall

$$\begin{aligned} T_n &= \frac{u(t_{n+1}) - u(t_n)}{h} - \Phi(t_n, u(t_n)) \\ &= \frac{u(t_{n+1}) - u(t_n)}{h} - c_1 f(t_n, u(t_n)) - c_2 f(t_n + a_2 h, u(t_n) + b_{21} f(t_n, u(t_n)) h) \end{aligned}$$

By expanding $u(t_n+h)$ about t_n, and $f(t_n+a_2 h, u(t_n)+b_{21} f(t_n, u(t_n) h))$ about t_n then $u(t_n)$, and noting that

$$\begin{aligned} u'(t_n) &= f(t_n, u(t_n)) \\ u''(t_n) &= \frac{d}{dt} f(t_n, u(t_n)) \\ &= \frac{\partial}{\partial t} f(t_n, u(t_n)) + \frac{\partial}{\partial u} f(t_n, u(t_n)) \frac{d}{dt} u(t_n) \\ &= \frac{\partial}{\partial t} f(t_n, u(t_n)) + f(t_n, u(t_n)) \frac{\partial}{\partial u} f(t_n, u(t_n)), \end{aligned}$$

we can show that

$$\begin{aligned} T_n = h^2 \Bigg(\frac{1}{6} u'''(t_n) - \frac{c_2}{2} \Bigg[& b_{21}^2 f(t_n, u(t_n)) \frac{\partial^2}{\partial u^2} f(t_n, u(t_n)) + \\ & a_2 b_{21} \frac{\partial^2}{\partial u \partial t} f(t_n, u(t_n)) + a_2^2 \frac{\partial^2}{\partial t^2} f(t_n, u(t_n)) \Bigg] \Bigg) + O(h^3) \end{aligned}$$

as long as we choose

$$1/2 = c_2 a_2 = c_2 b_{21}.$$