Numerical Analysis – Lecture 9[1]

3 Ordinary differential equations

Problem 3.1 We wish to approximate the exact solution of the *ordinary differential equation (ODE)*

$$\boldsymbol{y}' = \boldsymbol{f}(t, \boldsymbol{y}), \qquad t \geq 0, \tag{3.1}$$

where $\boldsymbol{y} \in \mathbb{R}^N$ and the function $\boldsymbol{f} : \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}^N$ is sufficiently 'nice'. (In principle, it is enough for $\boldsymbol{f}$ to be Lipschitz to ensure that the solution exists and is unique. Yet, for simplicity, we henceforth assume that $\boldsymbol{f}$ is analytic: in other words, we are always able to expand locally into Taylor series.) The equation (3.1) is accompanied by the initial condition $\boldsymbol{y}(0) = \boldsymbol{y}_0$.
Our purpose is to approximate $\boldsymbol{y}_{n+1} \approx \boldsymbol{y}(t_{n+1})$, $n = 0, 1, \ldots$, where $t_m = mh$ and the *time step* $h > 0$ is small, from $\boldsymbol{y}_0, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_n$ and equation (3.1).

Definition 3.2 A *one-step method* is a map $\boldsymbol{y}_{n+1} = \boldsymbol{\varphi}_h(t_n, \boldsymbol{y}_n)$, i.e. an algorithm which allows $\boldsymbol{y}_{n+1}$ to depend only on t_n, $\boldsymbol{y}_n$, h and the ODE (3.1).

Method 3.3 (Euler's method) We know $\boldsymbol{y}$ and its slope $\boldsymbol{y}'$ at $t = 0$ and wish to approximate $\boldsymbol{y}$ at $t = h > 0$. The most obvious approach is to truncate $\boldsymbol{y}(h) = \boldsymbol{y}(0) + h\boldsymbol{y}'(0) + \frac{1}{2}h^2\boldsymbol{y}''(0) + \cdots$ at the h^2 term. Since $\boldsymbol{y}'(0) = \boldsymbol{f}(t_0, \boldsymbol{y}_0)$, this procedure approximates $\boldsymbol{y}(h) \approx \boldsymbol{y}_0 + h\boldsymbol{f}(t_0, \boldsymbol{y}_0)$ and we thus set $\boldsymbol{y}_1 = \boldsymbol{y}_0 + h\boldsymbol{f}(t_0, \boldsymbol{y}_0)$.
By the same token, we may advance from h to $2h$ by letting $\boldsymbol{y}_2 = \boldsymbol{y}_1 + h\boldsymbol{f}(t_1, \boldsymbol{y}_1)$. In general, we obtain the *Euler method*

$$\boldsymbol{y}_{n+1} = \boldsymbol{y}_n + h\boldsymbol{f}(t_n, \boldsymbol{y}_n), \qquad n = 0, 1, \ldots. \tag{3.2}$$

Definition 3.4 Let $t^* > 0$ be given. We say that a method, which for every $h > 0$ produces the solution sequence $\boldsymbol{y}_n = \boldsymbol{y}_n(h)$, $n = 0, 1, \ldots, \lfloor t^*/h \rfloor$, *converges* if, as $h \to 0$ and $n_k(h)h \overset{k\to\infty}{\longrightarrow} t$, it is true that $\boldsymbol{y}_{n_k} \to \boldsymbol{y}(t)$, the exact solution of (3.1), uniformly for $t \in [0, t^*]$.

Theorem 3.5 *Suppose that $\boldsymbol{f}$ satisfies the Lipschitz condition: there exists $\lambda \geq 0$ such that*

$$\|\boldsymbol{f}(t, \boldsymbol{v}) - \boldsymbol{f}(t, \boldsymbol{w})\| \leq \lambda\|\boldsymbol{v} - \boldsymbol{w}\|, \qquad t \in [0, t^*], \quad \boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^N.$$

Then the Euler method (3.2) converges.

Proof Let $\boldsymbol{e}_n = \boldsymbol{y}_n - \boldsymbol{y}(t_n)$, the error at step n, where $0 \leq n \leq t^*/h$. Thus,

$$\boldsymbol{e}_{n+1} = \boldsymbol{y}_{n+1} - \boldsymbol{y}(t_{n+1}) = [\boldsymbol{y}_n + h\boldsymbol{f}(t_n, \boldsymbol{y}_n)] - [\boldsymbol{y}(t_n) + h\boldsymbol{y}'(t_n) + \mathcal{O}(h^2)].$$

By the Taylor theorem, the $\mathcal{O}(h^2)$ term can be bounded uniformly for all $[0, t^*]$ (in the underlying norm $\|\cdot\|$) by ch^2, where $c > 0$. Thus, using (3.1) and the triangle inequality,

$$\begin{aligned}\|\boldsymbol{e}_{n+1}\| &\leq \|\boldsymbol{y}_n - \boldsymbol{y}(t_n)\| + h\|\boldsymbol{f}(t_n, \boldsymbol{y}_n) - \boldsymbol{f}(t_n, \boldsymbol{y}(t_n))\| + ch^2 \\ &\leq \|\boldsymbol{y}_n - \boldsymbol{y}(t_n)\| + h\lambda\|\boldsymbol{y}_n - \boldsymbol{y}(t_n)\| + ch^2 = (1 + h\lambda)\|\boldsymbol{e}_n\| + ch^2.\end{aligned}$$

Consequently, by induction,

$$\|\boldsymbol{e}_{n+1}\| \leq (1 + h\lambda)^m\|\boldsymbol{e}_{n+1-m}\| + ch^2\sum_{j=0}^{m-1}(1 + h\lambda)^j, \qquad m = 0, 1, \ldots, n + 1.$$

[1]Please email all corrections and suggestions to these notes to A.Iserles@damtp.cam.ac.uk. All handouts are available on the WWW at the URL http://www.damtp.cam.ac.uk/user/na/PartII/Handouts.html.

In particular, letting $m = n+1$ and bearing in mind that $\boldsymbol{e}_0 = \mathbf{0}$, we have

$$\|\boldsymbol{e}_{n+1}\| \le ch^2 \sum_{j=0}^{n} (1+h\lambda)^j = ch^2 \frac{(1+h\lambda)^{n+1}-1}{(1+h\lambda)-1} \le \frac{ch}{\lambda}(1+h\lambda)^{n+1}.$$

But for small $h > 0$ it is true that $0 < 1+h\lambda \le \mathrm{e}^{h\lambda}$. This and $(n+1)h \le t^*$ imply that $(1+h\lambda)^{n+1} \le \mathrm{e}^{t^*\lambda}$, therefore

$$\|\boldsymbol{e}_n\| \le \frac{c\mathrm{e}^{t^*\lambda}}{\lambda} h \to 0, \qquad h \to 0, \quad \text{uniformly for} \quad 0 \le nh \le t^*$$

and the theorem is true. □

Definition 3.6 The *order* of a general numerical method $\boldsymbol{y}_{n+1} = \boldsymbol{\varphi}_h(t_n, \boldsymbol{y}_0, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_n)$ for the solution of (3.1) is the largest integer $p \ge 0$ such that

$$\boldsymbol{y}(t_{n+1}) - \boldsymbol{\varphi}_h(t_n, \boldsymbol{y}(t_0), \boldsymbol{y}(t_1), \ldots, \boldsymbol{y}(t_n)) = \mathcal{O}\big(h^{p+1}\big)$$

for all $h > 0$, $n \ge 0$ and all sufficiently smooth functions $\boldsymbol{f}$ in (3.1). Note that, unless $p \ge 1$, the 'method' is an unsuitable approximation to (3.1): in particular, $p \ge 1$ is necessary for convergence.

Remark 3.7 (The order of Euler's method) We now have $\boldsymbol{\varphi}_h(t, \boldsymbol{y}) = \boldsymbol{y} + h\boldsymbol{f}(t, \boldsymbol{y})$. Substituting the exact solution of (3.1), we obtain from the Taylor theorem

$$\boldsymbol{y}(t_{n+1}) - [\boldsymbol{y}(t_n) + h\boldsymbol{f}(t_n, \boldsymbol{y}(t_n))] = [\boldsymbol{y}(t_n) + h\boldsymbol{y}'(t_n) + \tfrac{1}{2}h^2\boldsymbol{y}''(t_n) + \cdots] - [\boldsymbol{y}(t_n) + h\boldsymbol{y}'(t_n)] = \mathcal{O}\big(h^2\big)$$

and we deduce that Euler's method is of order 1.

Definition 3.8 (Theta methods) We consider methods of the form

$$\boldsymbol{y}_{n+1} = \boldsymbol{y}_n + h[\theta \boldsymbol{f}(t_n, \boldsymbol{y}_n) + (1-\theta)\boldsymbol{f}(t_{n+1}, \boldsymbol{y}_{n+1})], \qquad n = 0, 1, \ldots, \tag{3.3}$$

where $\theta \in [0,1]$ is a parameter:

- If $\theta = 1$, we recover Euler's method.
- if $\theta \in [0,1)$ then the *theta method* (3.3) is *implicit:* Each time step requires the solution of N (in general, nonlinear) algebraic equations for the unknown vector $\boldsymbol{y}_{n+1}$.
- The choices $\theta = 0$ and $\theta = \frac{1}{2}$ are known as

$$\begin{aligned} \text{Backward Euler:} \quad & \boldsymbol{y}_{n+1} = \boldsymbol{y}_n + h\boldsymbol{f}(t_{n+1}, \boldsymbol{y}_{n+1}), \\ \text{Trapezoidal rule:} \quad & \boldsymbol{y}_{n+1} = \boldsymbol{y}_n + \tfrac{1}{2}h[\boldsymbol{f}(t_n, \boldsymbol{y}_n) + \boldsymbol{f}(t_{n+1}, \boldsymbol{y}_{n+1})]. \end{aligned}$$

Solution of nonlinear algebraic equations can be done by iteration. For example, for backward Euler, letting $\boldsymbol{y}_{n+1}^{[0]} = \boldsymbol{y}_n$, we may use

Direct iteration $\quad \boldsymbol{y}_{n+1}^{[j+1]} = \boldsymbol{y}_n + h\boldsymbol{f}(t_{n+1}, \boldsymbol{y}_{n+1}^{[j]});$

Newton–Raphson: $\quad \boldsymbol{y}_{n+1}^{[j+1]} = \boldsymbol{y}_{n+1}^{[j]} - \left[I - h\frac{\partial \boldsymbol{f}(t_{n+1}, \boldsymbol{y}_{n+1}^{[j]})}{\partial \boldsymbol{y}}\right]^{-1} [\boldsymbol{y}_{n+1}^{[j]} - \boldsymbol{y}_n - h\boldsymbol{f}(t_{n+1}, \boldsymbol{y}_{n+1}^{[j]})];$

Modified Newton–Raphson: $\quad \boldsymbol{y}_{n+1}^{[j+1]} = \boldsymbol{y}_{n+1}^{[j]} - \left[I - h\frac{\partial \boldsymbol{f}(t_n, \boldsymbol{y}_n)}{\partial \boldsymbol{y}}\right]^{-1} [\boldsymbol{y}_{n+1}^{[j]} - \boldsymbol{y}_n - h\boldsymbol{f}(t_{n+1}, \boldsymbol{y}_{n+1}^{[j]})]$

We will return to this topic later.

Remark 3.9 (The order of the theta method) It follows from (3.3) and Taylor's theorem that

$$\begin{aligned} &\boldsymbol{y}(t_{n+1}) - \boldsymbol{y}(t_n) - h[\theta\boldsymbol{y}'(t_n) + (1-\theta)\boldsymbol{y}'(t_{n+1})] \\ &= [\boldsymbol{y}(t_n) + h\boldsymbol{y}'(t_n) + \tfrac{1}{2}h^2\boldsymbol{y}''(t_n) + \tfrac{1}{6}h^3\boldsymbol{y}'''(t_n)] - \boldsymbol{y}(t_n) - \theta h\boldsymbol{y}'(t_n) \\ &\quad - (1-\theta)h[\boldsymbol{y}'(t_n) + h\boldsymbol{y}''(t_n) + \tfrac{1}{2}h^2\boldsymbol{y}'''(t_n)] + \mathcal{O}\big(h^4\big) \\ &= (\theta - \tfrac{1}{2})h^2\boldsymbol{y}''(t_n) + (\tfrac{1}{2}\theta - \tfrac{1}{3})h^3\boldsymbol{y}'''(t_n) + \mathcal{O}\big(h^4\big)\,. \end{aligned}$$

Therefore the theta method is of order 1, except that the trapezoidal rule is of order 2.