Numerical Solution of Differential Equations I

Endre Süli

Mathematical Institute
University of Oxford
2020

Lecture 4

General explicit one-step methods

Definition

An explicit *one-step method* is of the form

$$y_{n+1} = y_n + h\Phi(x_n, y_n; h, f), \qquad n = 0, 1, \dots, N-1; \quad y_0 = \text{given},$$

where Φ is a continuous function of its first three variables, such that the triplet

$$(\xi, \eta; h) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}_{>0}$$

and a function f, are mapped into an approximation

$$\eta + h\Phi(\xi, \eta; h, f) \in \mathbb{R}$$

to $y(\xi + h)$, the value at $x = \xi + h$ of the solution y to the I.V.P.

$$y'(x) = f(x, y(x)), \qquad y(\xi) = \eta.$$

General explicit one-step methods

Definition

An explicit *one-step method* is of the form

$$y_{n+1} = y_n + h\Phi(x_n, y_n; h, f), \qquad n = 0, 1, \ldots, N-1; \quad y_0 = \text{given},$$

where Φ is a continuous function of its first three variables, such that the triplet

$$(\xi, \eta; h) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}_{>0}$$

and a function f, are mapped into an approximation

$$\eta + h\Phi(\xi, \eta; h, f) \in \mathbb{R}$$

to $y(\xi + h)$, the value at $x = \xi + h$ of the solution y to the I.V.P.

$$y'(x) = f(x, y(x)), \qquad y(\xi) = \eta.$$

None: For simplicity, shall not indicate the dependence of Φ on f in what follows.

Runge–Kutta methods

Euler's method is only first-order accurate.

Runge–Kutta methods

Euler's method is only first-order accurate.

Runge–Kutta methods aim to achieve higher accuracy by sacrificing the efficiency of Euler's method through re-evaluating f at points intermediate between $(x_n, y(x_n))$ and $(x_{n+1}, y(x_{n+1}))$.

Runge–Kutta methods

Euler's method is only first-order accurate.

Runge–Kutta methods aim to achieve higher accuracy by sacrificing the efficiency of Euler's method through re-evaluating f at points intermediate between $(x_n, y(x_n))$ and $(x_{n+1}, y(x_{n+1}))$.

The general R-**stage Runge–Kutta family** is defined by

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

In compressed form, this information is usually displayed in the so-called Butcher tableau shown in the following figure:

$$
\begin{array}{c|c}
a = Be & B \\
\hline
 & c^{\mathsf{T}}
\end{array}
\qquad \text{where } e = (1, \ldots, 1)^{\mathsf{T}}.
$$

Figure: Butcher tableau of a Runge–Kutta method

One-stage Runge–Kutta methods

Suppose that $R = 1$. Then, the resulting one-stage Runge–Kutta method is simply Euler's explicit method:

$$y_{n+1} = y_n + hf(x_n, y_n).$$

Two-stage Runge–Kutta methods

Next, take $R = 2$, corresponding to the following family:

$$y_{n+1} = y_n + h(c_1 k_1 + c_2 k_2), \qquad (1)$$

where

$$k_1 = f(x_n, y_n), \qquad (2)$$
$$k_2 = f(x_n + a_2 h, y_n + b_{21} h k_1), \qquad (3)$$

and where the parameters c_1, c_2, a_2 and b_{21} are to be determined.

Two-stage Runge–Kutta methods

Next, take $R = 2$, corresponding to the following family:

$$y_{n+1} = y_n + h(c_1 k_1 + c_2 k_2), \tag{1}$$

where

$$\begin{aligned} k_1 &= f(x_n, y_n), & (2) \\ k_2 &= f(x_n + a_2 h, y_n + b_{21} h k_1), & (3) \end{aligned}$$

and where the parameters c_1, c_2, a_2 and b_{21} are to be determined.

Clearly (1)–(3) can be rewritten as a general one-step method, with

$$\Phi(x, y; h) := c_1 f(x, y) + c_2 f(x + a_2 h, y + b_{21} h f(x, y)).$$

By the consistency condition $\Phi(x, y; 0) = f(x, y)$, a method from this family will be consistent if, and only if,

$$c_1 + c_2 = 1.$$

By the consistency condition $\Phi(x, y; 0) = f(x, y)$, a method from this family will be consistent if, and only if,

$$c_1 + c_2 = 1.$$

Further conditions on the parameters are obtained by attempting to maximise the order of accuracy of the method.

By the consistency condition $\Phi(x, y; 0) = f(x, y)$, a method from this family will be consistent if, and only if,

$$c_1 + c_2 = 1.$$

Further conditions on the parameters are obtained by attempting to maximise the order of accuracy of the method.

By expanding the consistency error of (1)–(3) in powers of h:

$$T_n = \frac{1}{2}hy''(x_n) + \frac{1}{6}h^2y'''(x_n) - c_2h[a_2f_x + b_{21}f_yf]$$
$$- c_2h^2\left[\frac{1}{2}a_2^2f_{xx} + a_2b_{21}f_{xy}f + \frac{1}{2}b_{21}^2f_{yy}f^2\right] + \mathcal{O}(h^3).$$

Here we have used the abbreviations

$$f = f(x_n, y(x_n)), \qquad f_x = \frac{\partial f}{\partial x}(x_n, y(x_n)), \qquad etc.$$

By noting that $y'' = f_x + f_y f$, it follows that $T_n = \mathcal{O}(h^2)$ for any f provided that

$$a_2 c_2 = b_{21} c_2 = \frac{1}{2},$$

which implies that if $b_{21} = a_2$, $c_2 = 1/(2a_2)$ and $c_1 = 1 - 1/(2a_2)$ then the method is second-order accurate.

By noting that $y'' = f_x + f_y f$, it follows that $T_n = \mathcal{O}(h^2)$ for any f provided that

$$a_2 c_2 = b_{21} c_2 = \frac{1}{2},$$

which implies that if $b_{21} = a_2$, $c_2 = 1/(2a_2)$ and $c_1 = 1 - 1/(2a_2)$ then the method is second-order accurate.

This still leaves one free parameter, a_2, but it is easy to see that no choice of a_2 will make the method third-order accurate.

There are two well-known examples of second-order Runge–Kutta methods of the form (1)–(3):

a) **The modified Euler method:** In this case we take $a_2 = \frac{1}{2}$ to obtain

$$y_{n+1} = y_n + h\, f\left(x_n + \frac{1}{2}h, y_n + \frac{1}{2}hf(x_n, y_n)\right);$$

b) **The improved Euler method:** This is arrived at by choosing $a_2 = 1$ which gives

$$y_{n+1} = y_n + \frac{1}{2}h\left[f(x_n, y_n) + f(x_n + h, y_n + hf(x_n, y_n))\right].$$

For these two methods it is easily verified by Taylor series expansion that the consistency error is of the form, respectively,

$$
\begin{aligned}
T_n &= \frac{1}{6}h^2\left[f_y F_1 + \frac{1}{4}F_2\right] + \mathcal{O}(h^3),\\
T_n &= \frac{1}{6}h^2\left[f_y F_1 - \frac{1}{2}F_2\right] + \mathcal{O}(h^3),
\end{aligned}
$$

where

$$F_1 = f_x + ff_y \quad \text{and} \quad F_2 = f_{xx} + 2ff_{xy} + f^2 f_{yy}.$$

For these two methods it is easily verified by Taylor series expansion that the consistency error is of the form, respectively,

$$
\begin{aligned}
T_n &= \frac{1}{6}h^2\left[f_y F_1 + \frac{1}{4}F_2\right] + \mathcal{O}(h^3), \\
T_n &= \frac{1}{6}h^2\left[f_y F_1 - \frac{1}{2}F_2\right] + \mathcal{O}(h^3),
\end{aligned}
$$

where

$$
F_1 = f_x + ff_y \quad \text{and} \quad F_2 = f_{xx} + 2ff_{xy} + f^2 f_{yy}.
$$

The methods (1)–(3) are explicit two-stage Runge–Kutta methods.

Three-stage Runge–Kutta methods

Let us now suppose that $R = 3$ to illustrate the general idea. Thus, we consider the family of methods:

$$y_{n+1} = y_n + h\left[c_1 k_1 + c_2 k_2 + c_3 k_3\right],$$

where

$$\begin{aligned}
k_1 &= f(x, y), \\
k_2 &= f(x + ha_2, y + hb_{21}k_1), \\
k_3 &= f(x + ha_3, y + hb_{31}k_1 + hb_{32}k_2), \\
a_2 &= b_{21}, \quad a_3 = b_{31} + b_{32}.
\end{aligned}$$

Three-stage Runge–Kutta methods

Let us now suppose that $R = 3$ to illustrate the general idea. Thus, we consider the family of methods:

$$y_{n+1} = y_n + h\left[c_1 k_1 + c_2 k_2 + c_3 k_3\right],$$

where

$$\begin{aligned}
k_1 &= f(x, y), \\
k_2 &= f(x + ha_2, y + hb_{21}k_1), \\
k_3 &= f(x + ha_3, y + hb_{31}k_1 + hb_{32}k_2), \\
a_2 &= b_{21}, \quad a_3 = b_{31} + b_{32}.
\end{aligned}$$

Writing $b_{21} = a_2$ and $b_{31} = a_3 - b_{32}$ in the definitions of k_2 and k_3 respectively and expanding k_2 and k_3 into Taylor series about the point (x, y) yields:

$$
\begin{aligned}
k_2 &= f + ha_2(f_x + k_1 f_y) + \frac{1}{2}h^2 a_2^2(f_{xx} + 2k_1 f_{xy} + k_1^2 f_{yy}) + \mathcal{O}(h^3) \\
&= f + ha_2(f_x + ff_y) + \frac{1}{2}h^2 a_2^2(f_{xx} + 2ff_{xy} + f^2 f_{yy}) + \mathcal{O}(h^3) \\
&= f + ha_2 F_1 + \frac{1}{2}h^2 a_2^2 F_2 + \mathcal{O}(h^3),
\end{aligned}
$$

$$
\begin{aligned}
k_2 &= f + ha_2(f_x + k_1 f_y) + \frac{1}{2}h^2 a_2^2 (f_{xx} + 2k_1 f_{xy} + k_1^2 f_{yy}) + \mathcal{O}(h^3) \\
&= f + ha_2(f_x + ff_y) + \frac{1}{2}h^2 a_2^2 (f_{xx} + 2ff_{xy} + f^2 f_{yy}) + \mathcal{O}(h^3) \\
&= f + ha_2 F_1 + \frac{1}{2}h^2 a_2^2 F_2 + \mathcal{O}(h^3),
\end{aligned}
$$

where

$$
F_1 = f_x + ff_y \quad \text{and} \quad F_2 = f_{xx} + 2ff_{xy} + f^2 f_{yy},
$$

and

$$
\begin{aligned}
k_3 &= f + h\left\{a_3 f_x + \left[(a_3 - b_{32})k_1 + b_{32}k_2\right] f_y\right\} \\
&\quad + \frac{1}{2}h^2 \left\{a_3^2 f_{xx} + 2a_3\left[(a_3 - b_{32})k_1 + b_{32}k_2\right] f_{xy}\right. \\
&\quad \left. + \left[(a_3 - b_{32})k_1 + b_{32}k_2\right]^2 f_{yy}\right\} + \mathcal{O}(h^3) \\
&= f + ha_3 F_1 + h^2\left(a_2 b_{32} F_1 f_y + \frac{1}{2}a_3^2 F_2\right) + \mathcal{O}(h^3).
\end{aligned}
$$

Substituting these expressions for k_2 and k_3 gives

$$\begin{aligned}\Phi(x,y,h) &= (c_1+c_2+c_3)f + h(c_2a_2+c_3a_3)F_1 \\ &\quad + \frac{1}{2}h^2\left[2c_3a_2b_{32}F_1f_y + \left(c_2a_2^2+c_3a_3^2\right)F_2\right] + \mathcal{O}(h^3).\end{aligned}$$

Substituting these expressions for k_2 and k_3 gives

$$\begin{aligned}\Phi(x, y, h) &= (c_1 + c_2 + c_3)f + h(c_2 a_2 + c_3 a_3)F_1 \\ &+ \frac{1}{2}h^2\left[2c_3 a_2 b_{32} F_1 f_y + \left(c_2 a_2^2 + c_3 a_3^2\right) F_2\right] + \mathcal{O}(h^3).\end{aligned}$$

We match this with the Taylor series expansion:

$$\begin{aligned}\frac{y(x+h) - y(x)}{h} &= y'(x) + \frac{1}{2}hy''(x) + \frac{1}{6}h^2 y'''(x) + \mathcal{O}(h^3) \\ &= f + \frac{1}{2}hF_1 + \frac{1}{6}h^2\left(F_1 f_y + F_2\right) + \mathcal{O}(h^3).\end{aligned}$$

Substituting these expressions for k_2 and k_3 gives

$$\begin{aligned}\Phi(x,y,h) &= (c_1+c_2+c_3)f + h(c_2a_2+c_3a_3)F_1 \\ &\quad + \frac{1}{2}h^2\left[2c_3a_2b_{32}F_1f_y + \left(c_2a_2^2+c_3a_3^2\right)F_2\right] + \mathcal{O}(h^3).\end{aligned}$$

We match this with the Taylor series expansion:

$$\begin{aligned}\frac{y(x+h)-y(x)}{h} &= y'(x) + \frac{1}{2}hy''(x) + \frac{1}{6}h^2y'''(x) + \mathcal{O}(h^3) \\ &= f + \frac{1}{2}hF_1 + \frac{1}{6}h^2\left(F_1f_y + F_2\right) + \mathcal{O}(h^3).\end{aligned}$$

This yields:

$$\begin{aligned}c_1+c_2+c_3 &= 1, \\ c_2a_2+c_3a_3 &= \frac{1}{2}, \\ c_2a_2^2+c_3a_3^2 &= \frac{1}{3}, \\ c_3a_2b_{32} &= \frac{1}{6}.\end{aligned}$$

Solving this system of four equations for the six unknowns: $c_1, c_2, c_3, a_2, a_3, b_{32}$, we obtain a two-parameter family of 3-stage Runge–Kutta methods. We highlight two notable examples:

Solving this system of four equations for the six unknowns: $c_1, c_2, c_3, a_2, a_3, b_{32}$, we obtain a two-parameter family of 3-stage Runge–Kutta methods. We highlight two notable examples:

(i) **Heun's method** corresponds to

$$c_1 = \frac{1}{4}, \quad c_2 = 0, \quad c_3 = \frac{3}{4}, \quad a_2 = \frac{1}{3}, \quad a_3 = \frac{2}{3}, \quad b_{32} = \frac{2}{3},$$

yielding

$$\begin{aligned} y_{n+1} &= y_n + \frac{1}{4} h\left(k_1 + 3k_3\right), \\ k_1 &= f(x_n, y_n), \\ k_2 &= f\left(x_n + \frac{1}{3}h, y_n + \frac{1}{3}hk_1\right), \\ k_3 &= f\left(x_n + \frac{2}{3}h, y_n + \frac{2}{3}hk_2\right). \end{aligned}$$

(ii) **Standard third-order Runge–Kutta method**. This is arrived at by selecting

$$c_1 = \frac{1}{6}, \quad c_2 = \frac{2}{3}, \quad c_3 = \frac{1}{6}, \quad a_2 = \frac{1}{2}, \quad a_3 = 1, \quad b_{32} = 2,$$

yielding

$$\begin{aligned} y_{n+1} &= y_n + \frac{1}{6} h \left(k_1 + 4k_2 + k_3 \right), \\ k_1 &= f(x_n, y_n), \\ k_2 &= f\left(x_n + \frac{1}{2} h, y_n + \frac{1}{2} h k_1 \right), \\ k_3 &= f\left(x_n + h, y_n - h k_1 + 2 h k_2 \right). \end{aligned}$$

Four-stage Runge–Kutta methods

For $R = 4$, an analogous argument leads to a two-parameter family of four-stage Runge–Kutta methods of order four.

Four-stage Runge–Kutta methods

For $R = 4$, an analogous argument leads to a two-parameter family of four-stage Runge–Kutta methods of order four.

A particularly popular example from this family is:

$$y_{n+1} = y_n + \frac{1}{6}h\left(k_1 + 2k_2 + 2k_3 + k_4\right),$$

where

$$\begin{aligned}
k_1 &= f(x_n, y_n), \\
k_2 &= f\left(x_n + \frac{1}{2}h, y_n + \frac{1}{2}hk_1\right), \\
k_3 &= f\left(x_n + \frac{1}{2}h, y_n + \frac{1}{2}hk_2\right), \\
k_4 &= f(x_n + h, y_n + hk_3).
\end{aligned}$$

In this lecture, we have constructed R-stage Runge–Kutta methods of order of accuracy $\mathcal{O}(h^R)$, $R = 1, 2, 3, 4$.

In this lecture, we have constructed R-stage Runge–Kutta methods of order of accuracy $\mathcal{O}(h^R)$, $R = 1, 2, 3, 4$.

Question: Is there an R stage method of order R for $R \geq 5$?

In this lecture, we have constructed R-stage Runge–Kutta methods of order of accuracy $\mathcal{O}(h^R)$, $R = 1, 2, 3, 4$.

Question: Is there an R stage method of order R for $R \geq 5$?

The answer to this question is unfortunately negative: in a series of papers John Butcher showed that for $R = 5, 6, 7, 8, 9$, the highest order that can be attained by an R-stage Runge–Kutta method is, respectively, $4, 5, 6, 6, 7$, and that for $R \geq 10$ the highest order is $\leq R - 2$.