Chapter 2

Nonlinear One-step Methods

2.1 Methods with Higher Accuracy

The Euler method discussed earlier has global error $O(h)$. In this section we introduce some higher order methods.

Definition 2.1.1 *A numerical method of which the global error behaves like $O(h^r)$ for some $r > 0$ is called an r-th order method.*

Taylor's Series Method: The method is expressed as

$$y_{n+1} = y_n + hy_n^{'} + \frac{h^2}{2}y_n^{''} + \ldots + \frac{h^r}{r!}y_n^{(r)}, \tag{2.1}$$

where the higher order derivatives can be evaluated by using the chain rule. For example, we have

$$\begin{aligned} y_n^{'} &= f(x_n, y_n) \\ y^{''} &= f_x + f_y f \\ y^{'''} &= f_{xx} + 2f_{xy}f + f_x f_y + f_{yy}f^2 + f_y^2 f \\ &\vdots \end{aligned} \tag{2.2}$$

Evidently this is not a practical method unless the function f is simple enough. It should be noted that eventually all numerical methods can be reduced to the form (2.1) where the higher order derivatives are approximated by some approximate combinations of known values. Note that the local truncation error of (2.1) is $\frac{h^{r+1}y^{(r+1)}}{(r+1)!}$.

Richardson's Extrapolation Method: Let q be the quantity to be approximated and let $N(q)$ be the numerical approximation of q by a certain

numerical method. Usually, it will be such that

$$N(q) = q + Ch^m + O(h^n) \tag{2.3}$$

where $m \leq n$ and h is the step size.

Suppose we have used two step sizes $h = h_1$ and $h = h_2$ to obtain $N_1(q)$ and $N_2(q)$, respectively. Assume $h_2 \leq h_1$. Let $k := \frac{h_1}{h_2} > 1$. It is easy to see that

$$\begin{aligned} k^m N_2(q) - N_1(q) &= k^m q + ck^m h_2^m - q - ch_1^m + k^m O(h_2^n) - O(h_1^n) \\ &= (k^m - 1)q + O(h_1^n. \end{aligned}$$

We thus obtain

$$q = \frac{k^m N_2(q) - N_1(q)}{k^m - 1} + O(h_1^n). \tag{2.4}$$

That is, we have derived a higher order method (2.4) from two applications of (2.3).

Second-Order Runge Kutta Method: Applying the Euler method with step size h and $\frac{h}{2}$, respectively, we obtain

$$y_{n+1,h} = y_n + hf(x_n, y_n) \tag{2.5}$$

$$\begin{aligned} q_1 &= y_n + \frac{h}{2} f(x_n, y_n) \\ y_{n+1,\frac{h}{2}} &= q_1 + \frac{h}{2} f(x_n + \frac{h}{2}, q_1). \end{aligned} \tag{2.6}$$

From (2.4), we know

$$\begin{aligned} y_{n+1} &:= 2y_{n+1,\frac{h}{2}} - y_{n+1,h} + O(h^2) \\ &= y_n + hf(x_n + \frac{h}{2}, q_1) + O(h^2). \end{aligned} \tag{2.7}$$

The new scheme

$$q_1 := y_n + \frac{h}{2} f(x_n, y_n) \tag{2.8}$$

$$y_{n+1} := y_n + hf(x_n + \frac{h}{2}, q_1) \tag{2.9}$$

with global error $O(h^2)$ is called *the one-step midpoint method*. The function value $f(x_n + \frac{h}{2}, q_1)$ may be thought of as an approximation to the Taylor's series $hy_n^{'} + \frac{h^2}{2} y_n^{''} + \ldots$. Indeed, by (2.2), we have

$$\begin{aligned} f(x_n + \frac{h}{2}, q_1) &= f(x_n, y_n) + \frac{h}{2} \frac{\partial f(x_n, y_n)}{\partial x} + \frac{h}{2} \frac{\partial f(x_n, y_n)}{\partial y} + O(h^2) \\ &= y_n^{'} + \frac{h}{2} y_n^{''} + O(h^2). \end{aligned} \tag{2.10}$$

It follows that $y_{n+1} = y_n + hy_n^{'} + \frac{h^2}{2} y_n^{''} + O(h^3)$. The local truncation error is $O(h^3)$.

The scheme (2.9) suggests a more general approach:

$$\begin{aligned} q_1 &:= y_n + \alpha h f(x_n, y_n) && (2.11)\\ y_{n+1} &:= y_n + \beta h f(x_n, y_n) + \gamma h f(x_n + \eta h, , q_1) && (2.12) \end{aligned}$$

where $\alpha, \beta, \gamma, \eta$ are constants to be selected so that y_{n+1} in (2.12) agrees with the Taylor's series of $y(x_{n+1})$ as closely as possible. Upon substitution, it can be checked that

$$\begin{aligned} y_{n+1} \quad &= y_n + (\beta+\gamma)h y_n^{'} + \gamma\left(\eta\frac{\partial f}{\partial x} + \alpha\frac{\partial f}{\partial y}y_n^{'}\right)h^2 \\ &+\gamma\left(\frac{\alpha^2}{2}\frac{\partial^2 f}{\partial y^2}(y_n^{'})^2 + \alpha\eta\frac{\partial^2 f}{\partial x \partial y}y_n^{'} + \frac{1}{2}\eta^2\frac{\partial^2 f}{\partial x^2}\right)h^3 + O(h^4). \quad (2.13) \end{aligned}$$

It is therefore desirable to have

$$\begin{aligned} \beta + \gamma &= 1 && (2.14)\\ \alpha &= \eta = \frac{1}{2\gamma}. && (2.15) \end{aligned}$$

The methods defined by (2.11) and (2.12) with conditions (2.14) and (2.15) are called *the (1-parameter family) second-order Runge-Kutta methods.*

2.2 Explicit Runge-Kutta Methods

A general R-stage Runge-Kutta method is defined by the one-step method:

$$y_{n+1} = y_n + h\phi(x_n, y_n, h) \quad (2.16)$$

where

$$\begin{aligned} \phi(x_n, y_n, h) &= \sum_{r=1}^{R} c_r k_r && (2.17)\\ \sum_{r=1}^{R} c_r &= 1 \\ k_r &= f(x_n + a_r h, y_n + h\sum_{s=1}^{R} b_{rs}k_s) && (2.18)\\ \sum_{s=1}^{R} b_{rs} &= a_r. \end{aligned}$$

It is convenient to display the coefficients occurring in (2.17) and (2.18) in the following form, known as a *Butcher array*:

$$\begin{array}{c|cccc} a_1 & b_{11} & b_{12} & \dots & b_{1R} \\ a_2 & b_{21} & b_{22} & \dots & b_{2R} \\ \vdots & \vdots & & & \vdots \\ a_R & b_{R1} & b_{R2} & \dots & b_{RR} \\ \hline & c_1 & c_2 & \dots & c_R \end{array}$$

We define the R-dimensional vectors a and c and the $R \times R$ matrix B by

$$\begin{aligned} a &= [a_1, a_2, \dots, a_R]^T \\ c &= [c_1, c_2, \dots, c_R]^T \\ B &= [b_{ij}]. \end{aligned}$$

Then an R-stage Runge-Kutta method is completely specified by its Butcher array

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

Definition 2.2.1 *An R-stage Runge-Kutta method is said to be*

1. *Explicit if and only if B is strictly lower triangular.*
2. *Semi-Implicit if and only if B is lower triangular.*
3. *Implicit if and only if B is not lower triangular.*

Remark. Note that an R-stage Runge-Kutta method involves R function evaluations per step. Each of the functions k_r, $r = 1, \dots, R$ may be interpreted as an approximation to the derivative $y^{'}$ at x_n, and the function ϕ is a weighted mean of these approximations. The idea is to choose values for these coefficients a, B, c so that the expansion $\phi(x, y, h)$ agrees with the Taylor's series

$$\phi_T(x_n, y_n, h) = y_n^{'} + \frac{h}{2} y_n^{''} + \dots \tag{2.19}$$

as closely as possible. If ϕ differs from ϕ_T only in the r-th and higher powers of h, then the method (2.16) has order r.

Example. Consider an explicit 3-stage method. Write (2.19) as

$$\phi_T = f + \frac{1}{2} hF + \frac{1}{6} h^2 (Ff_y + G) + O(h^3) \tag{2.20}$$

with $F := f_x + f_y f$ and $G := f_{xx} + 2ff_{xy} + f^2 f_{yy}$. Upon expanding the function ϕ where

$$\begin{aligned} \phi &= c_1 k_1 + c_2 k_2 + c_3 k_3 \\ k_1 &= f(x_n, y_n) = f \end{aligned}$$

$$\begin{aligned} k_2 &= f(x_n + a_2 h, y_n + h a_2 k_1) \\ &= f + a_2 F h + \frac{1}{2} a_2^2 G h^2 + O(h^3) \\ k_3 &= f(x_n + a_3 h, y_n + h(a_3 - b_{32}) k_1 + h b_{32} k_2) \\ &= f + a_3 F h + (a_2 b_{32} F f_y + \frac{1}{2} a_3^2 G) h^2 + O(h^3), \end{aligned}$$

we obtain

$$\begin{aligned} \phi &= (c_1 + c_2 + c_3) f + (a_2 c_2 + a_3 c_3) F h \\ &\quad + \frac{1}{2} \left[2 a_2 b_{32} c_3 F f_y + (a_2^2 c_2 + a_3^2 c_3) G \right] h^2 + O(h^3). \end{aligned} \tag{2.21}$$

Comparing (2.20) with (2.21), we need the following equations to be satisfied:

$$\begin{aligned} c_1 + c_2 + c_3 &= 1 \\ a_2 c_2 + a_3 c_3 &= \frac{1}{2} \\ a_2^2 c_2 + a_3^2 c_3 &= \frac{1}{3} \\ a_2 b_{32} c_3 &= \frac{1}{6}. \end{aligned} \tag{2.22}$$

There are four equations in six unknowns. Thus there exist a 2-parameter family of solutions. Further consideration on the term of h^3 shows that no solution of (2.22) can eliminate the term of h^3. That is, there exist a doubly infinite family of 3-stage Runge-Kutta methods of order 3, and none of order higher than 3. Two particular solutions are

1. Huen's third-order formula:

0	0	0	0
$\frac{1}{3}$	$\frac{1}{3}$	0	0
$\frac{2}{3}$	0	$\frac{2}{3}$	0
	$\frac{1}{4}$	0	$\frac{3}{4}$

2. Kutta's third-order formula:

0	0	0	0
$\frac{1}{2}$	$\frac{1}{2}$	0	0
1	-1	2	0
	$\frac{1}{6}$	$\frac{2}{3}$	$\frac{1}{6}$

A great deal of tedious manipulation is involved when deriving Runge-Kutta methods of higher orders.

Example. There are 11 equations in 13 unknowns involed in an explicit 4-stage Runge-Kutta method. It can be shown that the resulting methods are of order 4 and no higher. The most well-known forth-order explicit Runge-Kutta method is given by

0	0	0	0	0
$\frac{1}{2}$	$\frac{1}{2}$	0	0	0
$\frac{1}{2}$	0	$\frac{1}{2}$	0	0
1	0	0	1	0
	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{6}$

Another forth-order Runge-Kutta method is given by

0	0	0	0	0
$\frac{1}{3}$	$\frac{1}{3}$	0	0	0
$\frac{2}{3}$	$-\frac{1}{3}$	1	0	0
1	1	-1	1	0
	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Theorem 2.2.1 *The attainable order $p^*(R)$ of an explicit R-stage Runge-Kutta method can be quantified as follows:*

$$p^*(R) \begin{cases} = R & R = 1, 2, 3, 4 \\ = R - 1 & R = 5, 6, 7 \\ = R - 2 & R = 8, 9 \\ \leq R - 2 & R \geq 10 \end{cases}$$

Proof. See, Math. Comput. 19(1965), pp. 408-417. □

2.3 Implicit Runge-Kutta Methods

In an implicit Runge-Kutta method the functions k_r are no longer explicitly defined. Instead, k_r must be determined through solving simultaneously a set of equations (2.1) which in general is nonlinear. The derivation of an implicit method is precisely the same as that of an explicit method, i.e., the coefficients are selected so that the expansion of ϕ agrees with ϕ_T as closely as possible.

Example. The method

$\frac{1}{2} + \frac{\sqrt{3}}{6}$	$\frac{1}{4}$	$\frac{1}{4} + \frac{\sqrt{3}}{6}$
$\frac{1}{2} - \frac{\sqrt{3}}{6}$	$\frac{1}{4} - \frac{\sqrt{3}}{6}$	$\frac{1}{4}$
	$\frac{1}{2}$	$\frac{1}{2}$

is the unique 2-stage implicit Runge-Kutta method of order 4.

Theorem 2.3.1 *For any $R \geq 2$, there exists an R-stage implicit Runge-Kutta method of order $2R$.*

Proof. See, J. Butcher, Coefficients for the study of Runge-Kutta integration process, J. Austral. Math. Soc., 3(1963), pp. 185-201. □

The considerable increase in attainable order does not necessarily give implicit methods more computational advantage because, at each step, an expensive nonlinear system of R equations needs to be solved. The use of implicit methods is almost exclusively restricted to stiff systems, in which context their superior stability properties justify the high cost of implementation.

2.4 Absolute Stability of Runge-Kutta Methods

Consider the test problem

$$y^{'} = \lambda y \tag{2.23}$$

with $\lambda \in C$. Applying a 3-stage explicit Runge-Kutta method to the problem, we should have

$$\begin{aligned} k_1 &= \lambda y_n, \\ k_2 &= \lambda y_n(1 + a_2\lambda h), \\ k_3 &= \lambda y_n(1 + a_3\lambda h + a_2 b_{32}\lambda^2 h^2). \end{aligned}$$

It follows that

$$y_{n+1} = y_n \left[1 + (c_1 + c_2 + c_3)\overline{h} + (a_2c_2 + a_3c_3)\overline{h}^2 + a_2b_{32}c_3\overline{h}^3\right] \tag{2.24}$$

with $\overline{h} := \lambda h$. Define

$$R(\overline{h}) := \sum_{n=0}^{3} \frac{\overline{h}^n}{n!}. \tag{2.25}$$

Observe that

$$e_{n+1} = e^{\overline{h}}y(x_n) - R(\overline{h})y_n = R(\overline{h})e_n + (e^{\overline{h}} - R(\overline{h}))y(x_n)$$

and that $R(\overline{h})$ agrees with $e^{\overline{h}}$ up to the h^3 term. We conclude that $R(\overline{h})$is the growth factor of the global error.

Definition 2.4.1 *The 3-stage explicit Runge-Kutta method is said to be absolutely stable on the domain* Γ *if* $R(\overline{h})$*, given by (2.25), satisfies* $|r| \leq 1$ *whenever* $\overline{h} \in \Gamma$.

Remark. It is important to note that *all* 3-state explicit Runge-Kutta methods of order 3 have the same growth factor.

The above result can be generalized as follows. Consider an R-stage explicit Runge-Kutta method of order R (This can happen only when $R \leq 4$.) By assumption, the value y_{n+1} given by the method differs from the Taylor expansion of the exact solution $y(x_{n+1})$ by terms of order h^{p+1}. Applying to THE TEST problem (2.23), it can be proved that

$$R(\overline{h}) = \sum_{n=0}^{R} \frac{\overline{h}^n}{n!}. \tag{2.26}$$

is the growth factor. For $R > 4$, the stability polynomial $R(\overline{h})$ gets more complicated.

Example. When λ is real, the region of absolute stability can be replaced by the interval of absolute stability. In this case, we have the following empirical observation for explicit Runge-Kutta methods:

R	IAS
1	(-2.00, 0)
2	(-2.00, 0)
3	(-2.51, 0)
4	(-2.78, 0)

From the table, we realize that if $\lambda << 0$, then it is necessary that $h << 1$. In other words, explicit Runge-Kutta methods are *not* suitable for solving stiff problem.

Example. Consider a 2-stage implicit Runge-Kutta method:

$$\begin{array}{c|cc} a_1 & b_{11} & b_{12} \\ a_2 & b_{21} & b_{22} \\ \hline & c_1 & c_2 \end{array}$$

Choose $c_1 = c_2 = \frac{1}{2}$, $a_1 = b_{11} = 0$, $a_2 = 1$ and $b_{21} = b_{22} = \frac{1}{2}$. Then we obtain a 2-stage implicit method of order 2, also known as the Trapezoidal rule:

$$y_{n+1} = y_n + \frac{h}{2}(y'_n + y'_{n+1}). \tag{2.27}$$

Applying (2.27) to the test problem, we obtain the finite difference equation

$$y_{n+1} = y_n + \frac{\overline{h}}{2}y_n + \frac{\overline{h}}{2}y_{n+1},$$

or equivalently,

$$y_{n+1} = \frac{1 + \frac{\overline{h}}{2}}{1 - \frac{\overline{h}}{2}}y_n. \tag{2.28}$$

The growth factor is thus

$$\gamma = \frac{1 + \frac{\overline{h}}{2}}{1 - \frac{\overline{h}}{2}} \tag{2.29}$$

If we write $\frac{\overline{h}}{2} = u + iv$, then

$$|\gamma| = |\frac{(1+u)+iv}{(1-u)-iv}| = |\frac{1+u^2+v^2+2u}{1+u^2+v^2-2u}|^{1/2}.$$

It follows that $|\gamma| < 1$ if and only if $u < 0$. In other words, the region of absolute stability contains the entire left-half complex plane.

Remark. As long as $\lambda < 0$, the step size h can be arbitrarily chosen to make $\overline{h}$ stay in the absolute stability region regardless the size of λ. Such a method is said to be *A-stable*.

Remark. The R-stage implicit Runge-Kutta method of order $2R$ is always A-stable. (See, B. L. Ehle, High order A-stable methods for the numerical solution of systems of DEs, BIT, 8(1968), pp. 276-278.)