CHAPTER 11

Numerical Ordinary Differential Equations

Consider a first order *ordinary differential equation* (ODE) of the form

$$y' = f(x, y),$$

where $y = y(x)$ is an unknown variable, $x \in \mathbb{R}$ is an independent variable and the function $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is given. Here, we used the notation $y' := dy/dx$. The objective is to find the solution y for all $x \in [a, b]$ subject to an *initial condition*

$$y(x_0) = y_0, \quad x_0 \in [a, b].$$

We call the problem of solving the above ODE along with this initial condition, the *initial value problem*.

It is well-known that there are many ODEs of physical interest that cannot be solved exactly although we know that such problems have unique solutions. If one wants the solution of such problems, then the only way is to obtain it approximately. One common way of obtaining an approximate solution to a given initial value problem is to numerically compute the solution using a numerical method (or numerical scheme). In this chapter, we introduce some basic numerical methods for approximating the solution of a given initial value problem.

In Section 11.1, we review the exact solvability of a given initial value problem and motivate the need of numerical methods. In Section 11.3, we introduce a basic numerical method called *Euler's method* for obtaining approximate solution of an initial value problem and discussed the error involved in this method. We also show in this section that the Euler method is of order 1. Modified forms of Euler method can be obtained

using numerical quadrature formulas as discussed in Section 11.4. Taylor approximation with higher order terms can be used to obtain numerical methods of higher order accuracy. However, such methods involve the higher order derivatives of the unknown function and are called *Taylor methods*. Taylor methods are hard to implement on a computer as they involve higher order derivatives. An alternate way of achieving higher order accuracy without using higher order derivatives of the unknown function is the famous *Runge-Kutta methods*. Runge-Kutta method of order 2 is derived in full detail in Section 11.5 and the formula for the Runge-Kutta method of order 4 is stated.

11.1 Review of Theory

In this section, we will discuss solving an initial value problem (IVP) for an ordinary differential equation. Let f be a continuous function of two variables defined on a domain $D \subset \mathbb{R}^2$. Let $(x_0, y_0) \in D$ be given. The initial value problem for an ODE is given by

$$y' = f(x, y), \qquad (x, y) \in D, \tag{11.1a}$$

$$y(x_0) = y_0. \tag{11.1b}$$

The following result is concerning the equivalence of the initial value problem (11.1) with an integral equation.

Lemma 11.1.1.

A continuous function y defined on an interval I containing the point x_0 is a solution of the initial value problem (11.1) if and only if y satisfies the integral equation

$$y(x) = y_0 + \int_{x_0}^{x} f(s, y(s))\, ds, \tag{11.2}$$

for all $x \in I$.

Proof.

If y is a solution of the initial value problem (11.1), then we have

$$y'(x) = f(x, y(x)). \tag{11.3}$$

Integrating the above equation from x_0 to x yields the integral equation (11.2).

On the other hand let y be a solution of integral equation (11.2). Observe that, due to continuity of the function $x \to y(x)$, the function $s \to f(s, y(s))$ is continuous

on I. Thus by fundamental theorem of integral calculus, the right hand side of (11.2) is a differentiable function with respect to x and its derivative is given by the function $x \to f(x, y(x))$, which is a continuous function. From (11.2), we see that y is continuously differentiable and hence a direct differentiation of this equation yields the ODE (11.1a). Evaluating (11.2) at $x = x_0$ gives the initial condition $y(x_0) = y_0$.

Remark 11.1.2.

The following are the consequences of Lemma 11.1.1.

- If f is a function of x only, ie., $f = f(x)$, which can be integrated explicitly, then the integration on the right hand side of (11.2) can be obtained explicitely to get the solution y of the IVP (11.1a) exactly.
  For instance, if $f(x, y) = e^x$, then the solution of (11.1a) is $y(x) = y_0 + e^x - e^{x_0}$.
- If f depends on y, ie., $f = f(x, y)$, then (11.2) is an integral equation which in general cannot be solved explicitly. However, there are some particular cases, for instance when (11.1a) is linear, separable or exact, then we can obtain a solution of (11.1a). There are certain other cases where an integral factor can be obtained through which the equation can be made exact. But, there are many ODEs for which none of the above mentioned methods can be used to obtain solution.
  For instance, the equation
  $$y' = x^2 + y^2$$
  is clearly neither linear nor separable nor exact. Also any of the standard ways of finding integrating factor will not work for this equation.
  When standard method cannot be used to obtain solutions exactly, several approximation procedures may be proposed for finding an approximate solution to the initial value problem (11.1). In these procedures, one may use the ODE (11.1a) itself or its equivalent integral formulation (11.2). If we choose the ODE (11.1a), then the derivative may be approximated by a numerical differentiation formula. In case we choose the integral formulation (11.2), then we may use various numerical integration techniques (quadrature rules) to adopt a numerical method for approximating the solution of the initial value problem (11.1).

Before going for a numerical method, it is important to ensure that the given initial value problem has a unique solution. Otherwise, we will be solving an initial value problem numerically, which actually does not have a solution or it may have many solutions and we do not know which solution the numerical method obtains. Let us

illustrate the non-uniqueness of solution of an initial value problem.

Example 11.1.3 [Peano].

Let us consider the initial value problem

$$y' = 3\,y^{2/3}, \quad y(0) = 0. \tag{11.4}$$

Note that $y(x) = 0$ for all $x \in \mathbb{R}$ is clearly a solution for this initial value problem. Also, $y(x) = x^3$ for all $x \in \mathbb{R}$. This initial value problem has infinite family of solutions parametrized by $c \geq 0$, given by

$$y_c(x) = \begin{cases} 0 & \text{if } x \leq c, \\ (x-c)^3 & \text{if } x > c, \end{cases}$$

defined for all $x \in \mathbb{R}$. Thus, we see that a solution to an initial value problem need not be unique.

However by placing additional conditions on f, we can achieve uniqueness as stated in the following theorem. The proof of this theorem is omitted for this course.

Theorem 11.1.4 [Cauchy-Lipschitz-Picard's Existence and Uniqueness Theorem].

Let $D \subseteq \mathbb{R}^2$ be a domain and $I \subset \mathbb{R}$ be an interval. Let $f : D \to \mathbb{R}$ be a continuous function. Let $(x_0, y_0) \in D$ be a point such that the rectangle R defined by

$$R = \{x \,:\, |x - x_0| \leq a\} \times \{y \,:\, |y - y_0| \leq b\} \tag{11.5}$$

is contained in D. Let f be Lipschitz continuous with respect to the variable y on R, *i.e.,* there exists a $K > 0$ such that

$$|f(x, y_1) - f(x, y_1)| \leq K|y_1 - y_2| \quad \forall (x, y_1),\, (x, y_2) \in R.$$

Then the initial value problem (11.1) has at least one solution on the interval $|x - x_0| \leq \delta$ where $\delta = \min\{a, \frac{b}{M}\}$. Moreover, the initial value problem (11.1) has exactly one solution on this interval.

Remark 11.1.5.

We state without proof that the function $f(x,y) = y^{2/3}$ in Example 11.1.3 is not a Lipschitz function on any rectangle containing the point $(0,0)$ in it.

While verifying that a given function f is Lipschitz continuous is a little difficult, one can easily give a set of alternative conditions that guarantee Lipschitz continuity, and also these conditions are easy to verify.

Lemma 11.1.6.

Let $D \subseteq \mathbb{R}^2$ be an open set and $f : D \to \mathbb{R}$ be a continuous function such that its partial derivative with respect to the variable y is also continuous on D, *i.e.,* $\frac{\partial f}{\partial y} : D \to \mathbb{R}$ is a continuous function. Let the rectangle R defined by

$$R = \{x \,:\, |x - x_0| \leq a\} \times \{y \,:\, |y - y_0| \leq b\}$$

be contained in D. Then f is Lipschitz continuous with respect to y in R.

Proof.

Let (x, y_1), $(x, y_2) \in R$. Applying mean value theorem with respect to the y variable, we get

$$f(x, y_1) - f(x, y_1) = \frac{\partial f}{\partial y}(x, \xi)(y_1 - y_2), \tag{11.6}$$

for some ξ between y_1 and y_2. Since we are applying mean value theorem by fixing x, this ξ will also depend on x. However since $\frac{\partial f}{\partial y} : D \to \mathbb{R}$ is a continuous function, it will be bounded on R. That is, there exists a number $L > 0$ such that

$$\left|\frac{\partial f}{\partial y}(x, y)\right| \leq L \quad \text{for all } (x, y) \in R. \tag{11.7}$$

Taking modulus in the equation (11.6), and using the last inequality we get

$$|f(x, y_1) - f(x, y_1)| = \left|\frac{\partial f}{\partial y}(x, \xi)\right| \, |(y_1 - y_2)| \leq L \, |(y_1 - y_2)|.$$

This finishes the proof of lemma.

The following theorem can be used as a tool to check the existence and uniqueness of solution of a given initial value problem (11.1).

Corollary 11.1.7 [Existence and Uniqueness Theorem].

Let $D \subseteq \mathbb{R}^2$ be a domain and $I \subseteq \mathbb{R}$ be an interval. Let $f : D \to \mathbb{R}$ be a continuous

function. Let $(x_0, y_0) \in D$ be a point such that the rectangle R defined by

$$R = \{x : |x - x_0| \leq a\} \times \{y : |y - y_0| \leq b\} \tag{11.8}$$

is contained in D.

If the partial derivative $\partial f/\partial y$ is also continuous in D, then there exists a unique solution $y = y(x)$ of the initial value problem (11.1) defined on the interval $|x - x_0| \leq \delta$ where $\delta = \min\{a, \frac{b}{M}\}$.

We state an example (without details) of an initial value problem that has a unique solution but the right hand side function in the differential equation is not a Lipschitz function. Thus this example illustrates the fact that condition of Lipschitz continuity is only a sufficient condition for uniqueness and by no means necessary.

Example 11.1.8.

The IVP

$$y' = \begin{cases} y \sin \frac{1}{y} & \text{if} \quad y \neq 0 \\ 0 & \text{if} \quad y = 0, \end{cases} \qquad y(0) = 0.$$

has unique solution, despite the RHS not being Lipschitz continuous with respect to the variable y on any rectangle containing $(0, 0)$.

Remark 11.1.9 [On Existence and Uniqueness Theorem].

1. Though we have stated existence theorems without proof, we can confirm that their proofs do not give an explicit soluton of the initial value problem being considered. In fact, the proofs give a sequence of functions that converge to a solution of the initial value problem.
2. Even when the RHS of the ordinary differential equation is a function of x only, we do not have explicit solutions. For example,

   $$y' = e^{-x^2}, \; y(0) = 0.$$

   The only alternative is to approximate its solution by a numerical procedure.

11.2 Discretization Notations

A numerical method gives approximate value of the solution of the initial value problem (11.1) at only a discrete set of point. Thus, if we are interested in obtaining solution

for (11.1a) in an interval $[a, b]$, then we first discretize the interval as

$$a = x_0 < x_1 < \cdots < x_n = b, \tag{11.9}$$

where each point x_i, $i = 0, 1, \cdots, n$ is called a *node*. Unless otherwise stated, we always assume that the nodes are equally spaced. That is,

$$x_j = x_0 + jh, \quad j = 0, 1, \cdots n \tag{11.10}$$

for a sufficiently small positive real number h. We use the notation for the approximate solution as

$$y_j = y_h(x_j) \approx y(x_j), \quad j = 0, 1, \cdots, n. \tag{11.11}$$

11.3 Euler's Method

The most simplest numerical method for a first order ordinary differential equation (11.1a) is obtained by replace the first order derivative of the unknown function by its finite difference formula. Assume that we know the value of the unknown function y at a point $x = x_0$. For obtaining the value of y at the point $x_1 = x_0 + h$, we use the forward difference formula for the derivative given by

$$y'(x) \approx \frac{1}{h}(y(x + h) - y(x))$$

in (11.1a) to get

$$\frac{1}{h}(y(x_1) - y(x_0)) \approx f(x_0, y(x_0)).$$

This can be used to obtain the value of $y(x_1)$ in terms of x_0 and $y(x_0)$ as

$$y(x_1) \approx y(x_0) + hf(x_0, y(x_0)).$$

Since we assumed that we know the value of $y(x_0)$, the right hand side is fully known and hence $y(x_1)$ can now be computed explicitly.

In general, if you know the value of $y(x_j)$, $j = 0, 1, \cdots, n$, we can obtain the value of $y(x_{j+1})$ by using the forward difference formula in (11.1a) at $x = x_j$ to get

$$\frac{1}{h}(y(x_{j+1}) - y(x_j)) \approx f(x_j, y(x_j)).$$

Denoting the approximate value of $y(x_j)$ by y_j, we can adopt the formula

$$y_{j+1} = y_j + hf(x_j, y_j), \quad j = 0, 1, \cdots, n - 1 \tag{11.12}$$

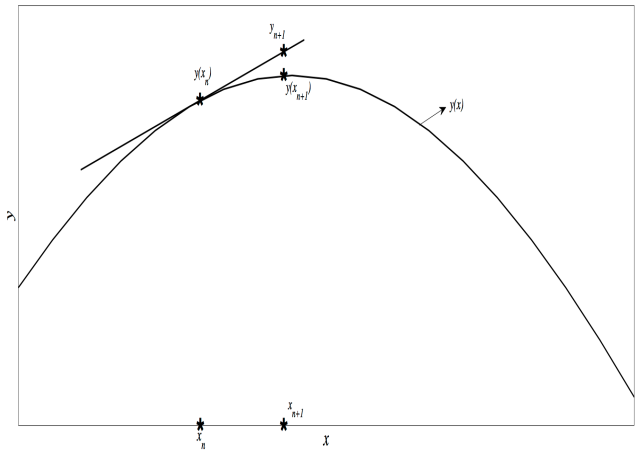

Figure 11.1: Geometrical interpretation of Euler's method

for obtaining the values of the solution y at the discrete points x_j, $j = 1, 2, \cdots, n$ by taking the value of y_0 from the initial condition (11.1b). The formula (11.12) is called the ***forward Euler's method***.

The ***backward Euler's method*** can be obtained by using the backward difference formula for the derivative of y in (11.1) and is given by

$$y_{j-1} = y_j - hf(x_j, y_j), \quad j = 0, -1, \cdots, -n+1. \tag{11.13}$$

By ***Euler's method*** we mean either forward or backward Euler's method depending on the context.

Remark 11.3.1 [Geometric Interpretation].

A geometric insight into Euler's method is shown in Figure 11.1 The tangent line to the graph of $y(x)$ at $x = x_j$ has slope $f(x_j, y_j)$. The Euler's method approximates the value of $y(x_{j+1})$ by the corresponding value of tangent line at the point $x = x_{j+1}$.

Example 11.3.2.

Consider the initial value problem

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

The Euler method (11.12) for this equation takes the form

$$y_{j+1} = y_j + hy_j = (1+h)y_j.$$

Note that the exact solution for the given initial value problem is $y(x) = e^x$.

On applying Euler's method with $h = 0.01$ and using 7-digit rounding, we get

$$\begin{aligned} y(0.01) &\approx y_1 = 1 + 0.01 = 1.01 \\ y(0.02) &\approx y_2 = 1.01 + 0.01(1.01) = 1.0201 \\ y(0.03) &\approx y_3 = 1.0201 + 0.01(1.0201) = 1.030301 \\ y(0.04) &\approx y_4 = 1.030301 + 0.01(1.030301) = 1.040604 \end{aligned}$$

The numerical results along with the error is presented in the following table for $h = 0.01$.

h	x	$y_h(x)$	Exact Solution	Error	Relative Error
0.01	0.00	1.000000	1.000000	0.000000	0.000000
0.01	0.01	1.010000	1.010050	0.000050	0.000050
0.01	0.02	1.020100	1.020201	0.000101	0.000099
0.01	0.03	1.030301	1.030455	0.000154	0.000149
0.01	0.04	1.040604	1.040811	0.000207	0.000199
0.01	0.05	1.051010	1.051271	0.000261	0.000248

Since the exact solution of this equation is $y = e^x$, the correct value at $x = 0.04$ is 1.040811.

By taking smaller values of h, we may improve the accuracy in the Euler's method. The numerical results along with the error is shown in the following table for $h = 0.005$.

h	x	$y_h(x)$	Exact Solution	Error	Relative Error
0.005	0.00	1.000000	1.000000	0.000000	0.000000
0.005	0.005	1.005000	1.005013	0.000013	0.000012
0.005	0.01	1.010025	1.010050	0.000025	0.000025
0.005	0.015	1.015075	1.015113	0.000038	0.000037
0.005	0.02	1.020151	1.020201	0.000051	0.000050
0.005	0.025	1.025251	1.025315	0.000064	0.000062
0.005	0.03	1.030378	1.030455	0.000077	0.000075
0.005	0.035	1.035529	1.035620	0.000090	0.000087
0.005	0.04	1.040707	1.040811	0.000104	0.000100
0.005	0.045	1.045910	1.046028	0.000117	0.000112
0.005	0.05	1.051140	1.051271	0.000131	0.000125

11.3.1 Error in Euler's Method

In Example (11.3.2), we illustrated that as we reduce the step size h, we tend to get more accurate solution of a given IVP at a given point $x = x_j$. The truncation error confirms this illustration when y'' is a bounded function. However, the mathematical error which involves the truncation error in the computed solution y_j and the propagating error from the computation of the solution at $x = x_i$ for $i = 0, 1, \cdots j-1$. In addition to the mathematical error, we also have arithmetic error due to floating-point approximation in each arithmetic operation. In this section, we study the total error involved in forward Euler's method. Total error involved in backward Euler's method can be obtained in a similar way.

Using Taylor's theorem, write

$$y(x_{j+1}) = y(x_j) + hy'(x_j) + \frac{h^2}{2}y''(\xi_j)$$

for some $x_j < \xi_j < x_{j+1}$. Since $y(x)$ satisfies the ODE $y' = f(x, y(x))$, we get

$$y(x_{j+1}) = y(x_j) + hf(x_j, y(x_j)) + \frac{h^2}{2}y''(\xi_j).$$

Thus, the local *truncation error* in forward Euler's method is

$$T_{j+1} = \frac{h^2}{2}y''(\xi_j), \tag{11.14}$$

which is the error involved in obtaining the value $y(x_{j+1})$ using the exact value $y(x_j)$. The forward Euler's method uses the approximate value y_j in the formula and therefore the finally computed value y_{j+1} not only involves the truncation error but also the

propagated error involved in computing y_j. Thus, the local *mathematical error* in the forward Euler's method is given by

$$\mathrm{ME}(y_{j+1}) := y(x_{j+1}) - y_{j+1} = y(x_j) - y_j + h(f(x_j, y(x_j)) - f(x_j, y_j)) + \frac{h^2}{2} y''(\xi_j).$$

Here, $y(x_j) - y_j + h(f(x_j, y(x_j)) - f(x_j, y_j))$ is the *propagated error*.

The propagated error can be simplified by applying the mean value theorem to $f(x, z)$ considering it as a functin of z:

$$f(x_j, y(x_j)) - f(x_j, y_j) = \frac{\partial f(x_j, \eta_j)}{\partial z}[y(x_j) - y_j],$$

for some η_j lying between $y(x_j)$ and y_j. Using this, we get the mathematical error

$$\mathrm{ME}(y_{j+1}) = \left[1 + h\frac{\partial f(x_j, \eta_j)}{\partial z}\right] \mathrm{ME}(y_j) + \frac{h^2}{2} y''(\xi_j) \tag{11.15}$$

for some $x_j < \xi_j < x_{j+1}$, and η_j lying between $y(x_j)$ and y_j.

We now assume that over the interval of interest,

$$\left|\frac{\partial f(x_j, y(x_j))}{\partial z}\right| < L, \quad |y''(x)| < Y,$$

where L and Y are fixed positive constants. On taking absolute values in (11.15), we obtain

$$|\mathrm{ME}(y_{j+1})| \leq (1 + hL)|\mathrm{ME}(y_j)| + \frac{h^2}{2} Y. \tag{11.16}$$

Applying the above estimate recursively, we get

$$\begin{aligned} |\mathrm{ME}(y_{j+1})| &\leq (1 + hL)^2 |\mathrm{ME}(y_{j-1})| + (1 + (1 + hL))\frac{h^2}{2} Y \\ &\leq \cdots \\ &\leq \cdots \\ &\leq (1 + hL)^{j+1} |\mathrm{ME}(y_0)| + \Big(1 + (1 + hL) + (1 + hL)^2 + \cdots + (1 + hL)^j\Big)\frac{h^2}{2} Y. \end{aligned}$$

Using the formulas

1. For any $\alpha \neq 1$,
$$1 + \alpha + \alpha^2 + \cdots + \alpha^j = \frac{\alpha^{j+1} - 1}{\alpha - 1}$$
2. For any $x \geq -1$,
$$(1 + x)^N \leq e^{Nx},$$

in the above inequality, we have proved the following theorem.

Theorem 11.3.3.

Let $y \in C^2[a,b]$ be a solution of the IVP (11.1) with

$$\left|\frac{\partial f(x,y)}{\partial y}\right| < L, \quad |y''(x)| < Y,$$

for all x and y, and some constants $L > 0$ and $Y > 0$. The mathematical error in the forward Euler's method at a point $x_j = x_0 + jh$ satisfies

$$|\mathrm{ME}(y_j)| \leq \frac{hY}{2L}\left(e^{(x_n - x_0)L} - 1\right) + e^{(x_n-x_0)L}|y(x_0) - y_0| \tag{11.17}$$

Example 11.3.4.

Consider the initial value problem

$$y' = y, \ y(0) = 1, \quad x \in [0,1].$$

Let us now find the upper bound for the mathematical error of forward Euler's method in solving this problem.

Here $f(x,y) = y$. Therefore, $\partial f/\partial y = 1$ and hence we can take $L = 1$.

Since $y = e^x$, $y'' = e^x$ and $|y''(x)| \leq e$ for $0 \leq x \leq 1$. Therefore, we take $Y = e$.

We now use the estimate (11.17) with $x_0 = 0$ and $x_n = 1$ to obtain

$$|\mathrm{ME}(y_j)| \leq \frac{he}{2}(e-1) \approx 2.3354h.$$

Here, we assume that there is no approximation in the initial condition and therefore the second term in (11.17) is zero.

To validate the upper bound obtained above, we shall compute the approximate solution of the given IVP using forward Euler's method. The method for the given IVP reads

$$y_{j+1} = y_j + hf(x_j, y_j) = (1+h)y_j.$$

The solution of this difference equation satisfing $y(0) = 1$ is

$$y_j = (1+h)^j.$$

Now, if $h = 0.1$, $n = 10$, we have $y_j = (1.1)^{10}$. Therefore, the forward Euler's method gives $y(1) \approx y_{10} \approx 2.5937$. But the exact value is $y(1) = e \approx 2.71828$. The error is 0.12466, whereas the bound obtained from (11.17) was 0.2354.

Remark 11.3.5.

The error bound (11.17) is valid for a large family of the initial value problems. But, it usually produces a very poor estimate due to the presence of the exponential terms. For instance, in the above example, if we take x_n to be very large, then the corresponding bound will also be very large.

The above error analysis assumes that the numbers used are of infinite precision and no floating point approximation is assumed. When we include the floating point approximation that $y_n = \tilde{y}_n + \epsilon_n$, then the bound for total error is given in the following theorem. The proof of this theorem is left as an exercise.

Theorem 11.3.6.

Let $y \in C^2[a, b]$ be a solution of the IVP (11.1) with

$$\left|\frac{\partial f(x, y)}{\partial y}\right| < L, \quad |y''(x)| < Y,$$

for all x and y, and some constants $L > 0$ and $Y > 0$. Let y_j be the approximate solution of (11.1) computed using the forward Euler's method (11.12) with infinite precision and let $\tilde{y}_j$ be the corresponding computed value using finite digit floating-point arithmetic. If

$$y_j = \tilde{y}_j + \epsilon_j,$$

then the total error $\text{TE}(y_j) := y(x_j) - \tilde{y}_j$ in forward Euler's method at a point $x_j = x_0 + jh$ satisfies

$$|\text{TE}(y_j)| \leq \frac{1}{L}\left(\frac{hY}{2} + \frac{\epsilon}{h}\right)\left(e^{(x_n - x_0)L} - 1\right) + e^{(x_n - x_0)L}|\epsilon_0|, \tag{11.18}$$

where $\epsilon := \max\{|\epsilon_i|/i = 0, 1, \cdots, n\}$.

11.4 Modified Euler's Methods

The Euler's method derived in the previous section can also be derived using the equivalent integral form of the IVP (11.1) as discussed in Lemma 11.1.1. Using this integral

form in the interval $[x_j, x_{j+1}]$, we get

$$y(x_{j+1}) = y(x_j) + \int_{x_j}^{x_{j+1}} f(s, y)ds. \tag{11.19}$$

The Euler's method can be obtained by replacing the integral on the right hand side by the rectangle rule.

Using the integral form in the interval $[x_{j-1}, x_{j+1}]$ and using the mid-point quadrature formula given by

$$\int_{x_{j-1}}^{x_{j+1}} f(s, y)ds \approx f(x_j, y_j)(x_{j+1} - x_{j-1}),$$

we get the ***Euler's mid-point method***

$$y_{j+1} = y_{j-1} + 2hf(x_j, y_j). \tag{11.20}$$

To compute the value of y_{j+1}, we need to know the value of y_{j-1} and y_j. Note that the above formula cannot be used to compute the value of y_1. Hence, we need another method to obtain y_1 and then y_j, for $j = 2, 3, \cdots, n$ can be obtained using (11.20). This method belongs to the class of ***2-step methods***.

Example 11.4.1.

Consider the initial-value problem

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

To obtain the approximate value of $y(0.04)$ with $h = 0.01$:

We first use Euler's method to get

$$y(0.01) \approx y_1 = 1 + 0.01 = 1.01$$

Next use Euler's mid-point method to get

$$\begin{aligned} y(0.02) &\approx y_2 = y_0 + 2 \times h \times y_1 = 1 + 2 \times 0.01 \times 1.01 = 1.0202 \\ y(0.03) &\approx y_3 = y_1 + 2 \times h \times y_2 = 1.01 + 2 \times 0.01 \times 1.0202 \approx 1.030404 \\ y(0.04) &\approx y_4 = y_2 + 2 \times h \times y_3 = 1.040808 \end{aligned}$$

Since the exact solution of this equation is $y = e^x$, the correct value at $x = 0.04$ is 1.040811. The error is 0.000003.

Recall the error in Euler method was 0.000199.

The methods derived above are ***explicit methods*** in the sense that the value of y_{j+1} is computed using the known values. If we using the trapezoidal rule for the integration on the right hand side of (11.19), we get

$$y_{j+1} = y_j + \frac{h}{2}(f(x_j, y_j) + f(x_{j+1}, y_{j+1})). \tag{11.21}$$

This method is called the ***Euler's Trapezoidal method***. Here, we see that the formula (11.21) involves an implicit relation for y_{j+1}. Such methods are referred to as ***implicit methods***.

Although the Euler's Trapezoidal method gives an implicit relation for y_{j+1}, sometimes it is explicit to compute the values y_{j+1} as illustrated in the following example.

Example 11.4.2.

Let us use the Euler's trapezoidal rule with $h = 0.2$ to obtain the appoximate solution of the initial value problem

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

We have $y_0 = 1$ and

$$y_1 = y_0 + \frac{h}{2}(x_0 y_0 + x_1 y_1) = 1 + 0.1(0 + 0.2y_1),$$

which gives $(1 - 0.02)y_1 = 1$, and this implies $y_1 \approx 1.0204$. Similarly,

$$y_2 = y_1 + \frac{h}{2}(x_1 y_1 + x_2 y_2) = 1.0204 + 0.1(0.2 \times 1.0204 + 0.4y_2),$$

and

$$y(0.4) \approx y_2 = \frac{1.0408}{1 - 0.04} \approx 1.0842.$$

In general, the Euler's trapezoidal rule gives a nonlinear equation for y_{j+1} as illustrated below.

Example 11.4.3.

Consider the initial value problem

$$y' = e^{-y}, \quad y(0) = 1.$$

We use the Euler's trapezoidal rule with $h = 0.2$ to solve the above problem. We

have,

$$y_1 = y_0 + \frac{h}{2}(e^{-y_0} + e^{-y_1}) = 1 + 0.1(e^{-1} + e^{-y_1}),$$

which gives the nonlinear equation

$$g(y_1) = y_1 - 0.1e^{-y_1} - (1 + 0.1e^{-1}) = 0,$$

and the solution of this equation is the approximate value of the solution $y(x_1)$ of the given initial value problem.

11.5 Runge-Kutta Methods

Although Euler's method is easy to implement, this method is not so efficient in the sense that to get a better approximation, one needs a very small step size. One way to get a better accuracy is to include the higher order terms in the Taylor expansion to get an approximation to y'. But the higher order terms involve higher derivatives of y. The ***Runge-Kutta methods*** attempts to obtain higher order accuracy and at the same time avoid the need for higher derivatives, by evaluating the function $f(x, y)$ at selected points on each subintervals. We first derive the Runge-Kutta method of order 2. The derivation of the Runge-Kutta method of order 4 can be done in a similar way. So, we skip the derivation of this method and present only final formula.

11.5.1 Order Two

Let y be a solution of the ODE (11.1a). The Runge-Kutta method of order 2 is obtained by truncating the Taylor expansion of $y(x + h)$ after the quadratic term. We derive now a formula for this method. Taylor expansion of $y(x + h)$ at the point x upto the quadratic term is given by

$$y(x + h) = y(x) + hy'(x) + \frac{h^2}{2}y''(x) + O(h^3). \tag{11.22}$$

Since y satisfies the given ODE $y' = f(x, y)$, by differentiating this ODE with respect to x both sides gives

$$y''(x) = \frac{\partial f}{\partial x}(x, y(x)) + y'(x)\frac{\partial f}{\partial y}(x, y(x)) \tag{11.23}$$

Substituting the values of y', y'' in (11.22), we get

$$y(x + h) = y(x) + hf(x, y(x)) + \frac{h^2}{2}\left[\frac{\partial f}{\partial x}(x, y(x)) + f(x, y(x))\frac{\partial f}{\partial y}(x, y(x))\right] + O(h^3).$$

The last equation is re-written as

$$y(x+h) = y(x) + \frac{h}{2} f(x, y(x)) + \frac{h}{2}\left[f(x, y(x)) + h\frac{\partial f}{\partial x}(x, y(x)) + hf(x, y(x))\frac{\partial f}{\partial y}(x, y(x)) \right] + O(h^3) \quad (11.24)$$

Taking $x = x_j$ for $j = 0, 1, \cdots, n-1$ with $x_{j+1} = x_j + h$ in (11.24), we get

$$y(x_{j+1}) = y(x_j) + \frac{h}{2} f(x_j, y(x_j)) + \frac{h}{2}\left[f(x_j, y(x_j)) + h\frac{\partial f}{\partial x}(x_j, y(x_j)) + hf(x_j, y(x_j))\frac{\partial f}{\partial y}(x_j, y(x_j)) \right] + O(h^3). \quad (11.25)$$

Let us now expand the function $f = f(s,t)$, which is a function of two variables, into its Taylor series at the point (ξ, τ) and truncate the series after the linear term. It is given by

$$f(s,t) = f(\xi,\tau) + (s-\xi)\frac{\partial f}{\partial s}(\xi,\tau) + (t-\tau)\frac{\partial f}{\partial t}(\xi,\tau) + O\big((s-\xi)^2\big) + O\big((t-\tau)^2\big).$$

Taking $(\xi, \tau) = (x_j, y(x_j))$ and comparing the above equation with the term in the square brackets in the equation (11.25), we get

$$y(x_{j+1}) = y(x_j) + \frac{h}{2} f(x_j, y(x_j)) + \frac{h}{2}\left[f(x_{j+1}, y(x_j) + hf(x_j, y(x_j))) \right] + O(h^3). \quad (11.26)$$

Truncating the higher order terms and denoting the approximate value of $y(x_{j+1})$ as y_{j+1}, we get

$$y_{j+1} = y_j + \frac{h}{2} f(x_j, y_j) + \frac{h}{2}\left[f(x_{j+1}, y_j + hf(x_j, y_j)) \right]. \quad (11.27)$$

Although the terms dropped from (11.26) to get (11.27) are of order 3 (namely, $O(h^3)$), the resultant approximation to y' is of order 2 as is evident from the Taylor's formula (11.22). The equation (11.27) is therefore known as ***Runge-Kutta method of order 2***. The formula (11.27) may be written as

$$y_{j+1} = y_j + \frac{h}{2}(k_1 + k_2),$$

where

$$\begin{aligned} k_1 &= f(x_j, y_j) \\ k_2 &= f(x_{j+1}, y_j + hk_1). \end{aligned}$$

The truncation error of Runge-Kutta method of order 2 is of order $O(h^3)$ whereas the Euler's method is of order $O(h^2)$. Therefore, for a fixed $h > 0$ we expect to get more accurate result from Runge-Kutta method order 2 when compared to Euler's method.

Example 11.5.1.

Consider the initial-value problem

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

Using Runge-Kutta method of order 2, we obtain

x	y	k_1	k_2
0.000000	1.000000	0.010000	0.010100
0.010000	1.010050	0.010000	0.010100
0.020000	1.020201	0.010100	0.010202
0.030000	1.030454	0.010202	0.010304
0.040000	1.040810	0.010305	0.010408

Recall the exact solution is $y(x) = e^x$ and $y(0.04) \approx 1.040811$. Therefore, the error involved is 0.000001 which is much less than the error (0.000199) obtained in Euler's method for $h = 0.01$.

11.5.2 Order Four

We state without derivation, the formula for the ***Runge-Kutta method of order 4***.

$$y_{j+1} = y_j + \frac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4)$$

where

$$\begin{aligned} k_1 &= f(x_j, y_j), \\ k_2 &= f\left(x_j + \frac{h}{2}, y_j + \frac{h}{2}k_1\right), \\ k_3 &= f\left(x_j + \frac{h}{2}, y_j + \frac{h}{2}k_2\right), \\ k_4 &= f(x_j + h, y_j + hk_3) \end{aligned}$$

The local truncation error of the 4^{th} order Runge-Kutta Method is of $O(h^5)$.

Example 11.5.2.

Consider the initial-value problem

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

Using Runge-Kutta method of order 4, we obtain

x_j	y_j	Exact Solution	Error	Relative Error
0.00	1.000000	1.000000	0.000000	0.000000
0.01	1.010050	1.010050	0.000000	0.000000
0.02	1.020201	1.020201	0.000000	0.000000
0.03	1.030455	1.030455	0.000000	0.000000
0.04	1.040811	1.040811	0.000000	0.000000

Note that the exact solution is $y(x) = e^x$.

11.6 Exercises

1. Let $h > 0$ and let $x_j = x_0 + jh$ $(j = 1, 2, \cdots, n)$ be given nodes. Consider the initial value problem $y'(x) = f(x, y), \quad y(x_0) = y_0$, with

$$\frac{\partial f(x,y)}{\partial y} \leq 0,$$

for all $x \in [x_0, x_n]$ and for all y.

   i) Using error analysis of the Euler's method, show that there exists an $h > 0$ such that

$$|e_n| \leq |e_{n-1}| + \frac{h^2}{2} f''(\xi) \quad \text{for some } \xi \in (x_{n-1}, x_n),$$

   where $e_n = y(x_n) - y_n$ with y_n obtained using Euler method.

   ii) Applying the conclusion of (i) above recursively, prove that

$$|e_n| \leq |e_0| + n\, h^2\, Y \quad \text{where } Y = \frac{1}{2} \max_{x_0 \leq x \leq x_n} |y''(x)|. \qquad (**)$$

2. The solution of the initial value problem

$$y'(x) = \lambda y(x) + \cos x - \lambda \sin x, \quad y(0) = 0$$

is $y(x) = \sin x$. For $\lambda = -20$, find the approximate value of $y(3)$ using the Euler's method with $h = 0.5$. Compute the error bound given in (), and Show that the actual absolute error exceeds the computed error bound given in (). Explain why it does not contradict the validity of ().

3. Derive the backward Euler's method for finding the approximate value of $y(x_n)$ for some $x_n < 0$, where y satisfies the initial value problem $y'(x) = f(x, y),\ \ y(0) = y_0$.

4. Consider the initial value problem $y' = -2y,\ 0 \le x \le 1,\ y(0) = 1$.
   i) Find an upper bound on the error in approximating the value of $y(1)$ computed using the Euler's method (at $x = 1$) in terms of the step size h.
   ii) For each h, solve the difference equation which results from the Euler's method, and obtain an approximate value of $y(1)$.
   iii) Find the error involved in the approximate value of $y(1)$ obtained in (ii) above by comparing with the exact solution.
   iv) Compare the error bound obtained in (i) with the actual error obtained in (iii) for $h = 0.1$, and for $h = 0.01$.
   v) If we want the absolute value of the error obtained in (iii) to be at most 0.5×10^{-6}, then how small the step size h should be?

5. Consider the initial value problem $y' = xy,\ y(0) = 1$. Estimate the error involved in the approximate value of $y(1)$ computed using the Euler's method (with infinite precision) with step size $h = 0.01$.

6. Find an upper bound for the propagated error in Euler method (with infinite precision) with $h = 0.1$ for solving the initial value problem $y' = y,\ \ y(0) = 1$, in the interval
   i) $[0, 1]$ and
   ii) $[0, 5]$.

7. For each $n \in \mathbb{N}$, write down the Euler's method for the solution of the initial value problem $y' = y,\ \ y(0) = 1$ on the interval $[0, 1]$ with step size $h = 1/n$. Let the resulting approximation to $y(1)$ be denoted by α_n. Show using limiting argument (without using the error bound) that $\alpha_n \to y(1)$ as $n \to \infty$.

8. In each of the following initial value problems, use the Euler's method, Runge-Kutta method of order 2 and 4 to find the solution at the specified point with specified step size h:
   i) $y' = x + y;\ \ y(0) = 1$. Find $y(0.2)$ (For the Euler's method take $h = 0.1$ and for other methods, take $h = 0.2$). The exact solution is $y(x) = -1 - x + 2e^x$.
   ii) $y' = 2\cos x - y,\ y(0) = 1$. Find $y(0.6)$ (For the Euler's method take $h = 0.1$ and for other methods, take $h = 0.2$) The exact solution is $y(x) = \sin x + \cos x$.

9. Use the Euler's, Runge-Kutta methods of order 2 and 4 to solve the IVP $y' = 0.5(x - y)$ for all $x \in [0, 3]$ with initial condition $y(0) = 1$. Compare the solutions for $h = 1, 0.5, 0.25, 0.125$ along with the exact solution $y(x) = 3e^{-x/2} + x - 2$.

10. Show that the Euler's and Runge-Kutta methods fail to determine an approximation to the non-trivial solution of the initial value problem $y' = y^\alpha,\ 0 < \alpha < 1$, $y(0) = 0$. Note that the ODE is in separable form and hence a non-trivial solution to initial value problem can be found analytically.

11. Write the formula using the Euler's method for approximating the solution to the initial value problem
$$y' = x^2 + y^2, \quad y(x_0) = y_0$$
at the point $x = x_1$ with $h = x_0 - x_1 > 0$. Find the approximate value y_1 of the solution to this initial value problem at the point x_1 when $x_0 = 0$, $y_0 = 1$, and $h = 0.25$. Find a bound for the truncation error in obtaining y_1.

12. Using the Gaussian rule determine a formula for solving the initial value problem
$$y' = e^{-x^2}, \quad y(x_0) = y_0.$$
in the form
$$y_{j+1} = y_{j-1} + h\left(e^{-(x_j - \frac{h}{\sqrt{3}})^2} + e^{-(x_j + \frac{h}{\sqrt{3}})^2}\right)$$
when the nodes are **equally spaced** with spacing $h = x_{j+1} - x_j$, $j \in \mathbb{Z}$. $(h > 0)$. Let $x_0 = 0$, and $y_0 = 1$. Using the method derived above, obtain approximate values of $y(-0.1)$ and $y(0.1)$.

Index