13

Numerical methods

Very few of the mathematical problems which arise in physical sciences and engineering can be solved analytically. Therefore, a simple, perhaps crude, technique giving the desired values within specified limits of tolerance is often to be preferred. We do not give a full coverage of numerical analysis in this chapter; but some methods for numerically carrying out the processes of interpolation, finding roots of equations, integration, and solving ordinary differential equations will be presented.

Interpolation

In the eighteenth century Euler was probably the first person to use the interpolation technique to construct planetary elliptical orbits from a set of observed positions of the planets. We discuss here one of the most common interpolation techniques: the polynomial interpolation. Suppose we have a set of observed or measured data (x_0, y_0), $(x_1, y_1), \ldots, (x_n, y_n)$, how do we represent them by a smooth curve of the form $y = f(x)$? For analytical convenience, this smooth curve is usually assumed to be polynomial:

$$f(x) = a_0 + a_1 x^1 + a_2 x^2 + \cdots + a_n x^n \tag{13.1}$$

and we use the given points to evaluate the coefficients $a_0, a_1, \ldots, a_n$:

$$\left.\begin{aligned} f(x_0) &= a_0 + a_1 x_0 + a_2 x_0^2 + \cdots + a_n x_0^n = y_0, \\ f(x_1) &= a_0 + a_1 x_1 + a_2 x_1^2 + \cdots + a_n x_1^n = y_1, \\ &\vdots \\ f(x_n) &= a_0 + a_1 x_n + a_2 x_n^2 + \cdots + a_n x_n^n = y_n. \end{aligned}\right\} \tag{13.2}$$

This provides $n+1$ equations to solve for the $n+1$ coefficients $a_0, a_1, \ldots, a_n$. However, straightforward evaluation of coefficients in the way outlined above is rather tedious, as shown in Problem 13.1, hence many shortcuts have been devised, though we will not discuss these here because of limited space.

Finding roots of equations

A solution of an equation $f(x) = 0$ is sometimes called a root, where $f(x)$ is a real continuous function. If $f(x)$ is sufficiently complicated that a direct solution may not be possible, we can seek approximate solutions. In this section we will sketch some simple methods for determining the approximate solutions of algebraic and transcendental equations. A polynomial equation is an algebraic equation. An equation that is not reducible to an algebraic equation is called transcendental. Thus, $\tan x - x = 0$ and $e^x + 2\cos x = 0$ are transcendental equations.

Graphical methods

The approximate solution of the equation

$$f(x) = 0 \tag{13.3}$$

can be found by graphing the function $y = f(x)$ and reading from the graph the values of x for which $y = 0$. The graphing procedure can often be simplified by first rewriting Eq. (13.3) in the form

$$g(x) = h(x) \tag{13.4}$$

and then graphing $y = g(x)$ and $y = h(x)$. The x values of the intersection points of the two curves gives the approximate values of the roots of Eq. (13.4). As an example, consider the equation

$$f(x) = x^3 - 146.25x - 682.5 = 0,$$

we can graph

$$y = x^3 - 146.25x - 682.5$$

to find its roots. But it is simpler to graph the two curves

$$y = x^3 \text{ (a cubic)}$$

and

$$y = 146.25x + 682.5 \text{ (a straight line)}.$$

See Fig. 13.1.

There is one drawback of graphical methods: that is, they require plotting curves on a large scale to obtain a high degree of accuracy. To avoid this, methods of successive approximations (or simple iterative methods) have been devised, and we shall sketch a couple of these in the following sections.

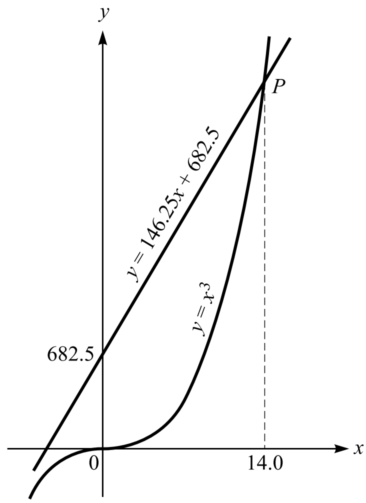

Figure 13.1.

Method of linear interpolation (method of false position)

Make an initial guess of the root of Eq. (13.3), say x_0, located between x_1 and x_2, and in the interval (x_1, x_2) the graph of $y = f(x)$ has the appearance as shown in Fig. 13.2. The straight line connecting P_1 and P_2 cuts the x-axis at point x_3, which is usually closer to x_0 than either x_1 or x_2. From similar triangles

$$\frac{x_3 - x_1}{-f(x_1)} = \frac{x_2 - x_1}{f(x_2)},$$

and solving for x_3 we get

$$x_3 = \frac{x_1 f(x_2) - x_2 f(x_1)}{f(x_2) - f(x_1)}.$$

Now the straight line connecting the points P_3 and P_2 intersects the x-axis at point x_4, which is a closer approximation to x_0 than x_3. By repeating this process we obtain a sequence of values $x_3, x_4, \ldots, x_n$ that generally converges to the root of the equation.

The iterative method described above can be simplified if we rewrite Eq. (13.3) in the form of Eq. (13.4). If the roots of

$$g(x) = c \tag{13.5}$$

can be determined for every real c, then we can start the iterative process as follows. Let x_1 be an approximate value of the root x_0 of Eq. (13.3) (and, of

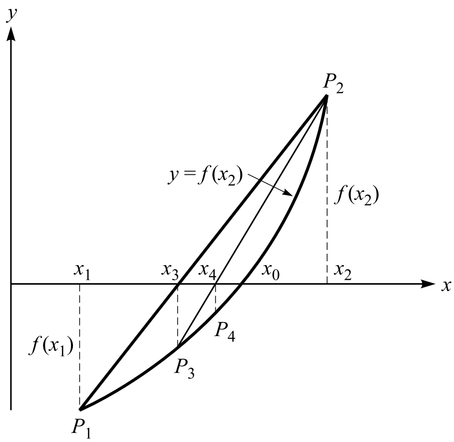

Figure 13.2.

course, also equation 13.4). Now setting $x = x_1$ on the right hand side of Eq. (13.4) we obtain the equation

$$g(x) = h(x_1), \tag{13.6}$$

which by hypothesis we can solve. If the solution is x_2, we set $x = x_2$ on the right hand side of Eq. (13.4) and obtain

$$g(x) = h(x_2). \tag{13.7}$$

By repeating this process, we obtain the nth approximation

$$g(x) = h(x_{n-1}). \tag{13.8}$$

From geometric considerations or interpretation of this procedure, we can see that the sequence $x_1, x_2, \ldots, x_n$ converges to the root $x = 0$ if, in the interval $2|x_1 - x_0|$ centered at x_0, the following conditions are met:

$$\left.\begin{array}{l} (1)\ \ |g'(x)| > |h'(x)|, \text{ and} \\ (2)\ \ \text{The derivatives are bounded.} \end{array}\right\} \tag{13.9}$$

Example 13.1

Find the approximate values of the real roots of the transcendental equation

$$e^x - 4x = 0.$$

Solution: Let $g(x) = x$ and $h(x) = e^x/4$, so the original equation can be rewritten as

$$x = e^x/4.$$

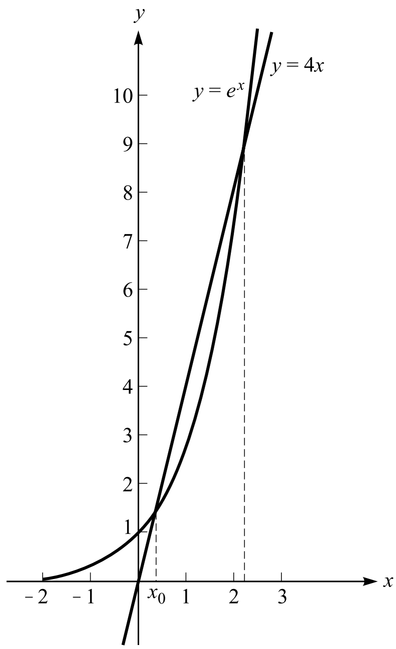

Figure 13.3.

According to Eq. (13.8) we have

$$x_{n+1} = e^{x_n}/4, \quad n = 1, 2, 3, \ldots. \tag{13.10}$$

There are two roots (see Fig. 13.3), with one around $x = 0.3$. If we take it as x_1, then we have, from Eq. (13.10)

$$x_2 = e^{x_1}/4 = 0.3374,$$
$$x_3 = e^{x_2}/4 = 0.3503,$$
$$x_4 = e^{x_3}/4 = 0.3540,$$
$$x_5 = e^{x_4}/4 = 0.3565,$$
$$x_6 = e^{x_5}/4 = 0.3571,$$
$$x_7 = e^{x_6}/4 = 0.3573.$$

The computations can be terminated at this point if only three-decimal-place accuracy is required.

The second root lies between 2 and 3. As the slope of $y = 4x$ is less than that of $y = e^x$, the first condition of Eq. (13.9) cannot be met, so we rewrite the original equation in the form

$$e^x = 4x, \quad \text{or} \quad x = \log 4x$$

and take $g(x) = x$, $h(x) = \log 4x$. We now have

$$x_{n+1} = \log 4x_n, \quad n = 1, 2, \ldots.$$

If we take $x_1 = 2.1$, then

$$x_2 = \log 4x_1 = 2.12823,$$
$$x_3 = \log 4x_2 = 2.14158,$$
$$x_4 = \log 4x_3 = 2.14783,$$
$$x_5 = \log 4x_4 = 2.15075,$$
$$x_6 = \log 4x_5 = 2.15211,$$
$$x_7 = \log 4x_6 = 2.15303,$$
$$x_8 = \log 4x_7 = 2.15316,$$

and we see that the value of the root correct to three decimal places is 2.153.

Newton's method

In Newton's method, the successive terms in the sequence of approximate values $x_1, x_2, \ldots, x_n$ that converges to the root is obtained by the intersection with the x-axis of the tangent line to the curve $y = f(x)$. Fig. 13.4 shows a portion of the graph of $f(x)$ close to one of its roots, x_0. We start with x_1, an initial guess of the value of the root x_0. Now the equation of the tangent line to $y = f(x)$ at P_1 is

$$y - f(x_1) f'(x_1)(x - x_1). \tag{13.11}$$

This tangent line intersects the x-axis at x_2 that is a better approximation to the root than x_1. To find x_2, we set $y = 0$ in Eq. (13.11) and find

$$x_2 = x_1 - f(x_1)/f'(x_1)$$

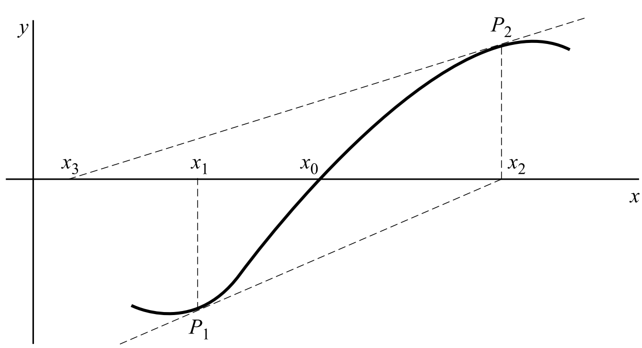

Figure 13.4.

https://doi.org/10.1017/CBO9780511755781.014 Published online by Cambridge University Press

provided $f'(x_1) \neq 0$. The equation of the tangent line at P_2 is

$$y - f(x_2) = f'(x_2)(x - x_2)$$

and it intersects the x-axis at x_3:

$$x_3 = x_2 - f(x_2)/f'(x_2).$$

This process is continued until we reach the desired level of accuracy. Thus, in general

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad n = 1, 2, \ldots. \tag{13.12}$$

Newton's method may fail if the function has a point of inflection, or other bad behavior, near the root. To illustrate Newton's method, let us consider the following trivial example.

Example 13.2
Solve, by Newton's method, $x^3 - 2 = 0$.

Solution: Here we have $y = x^3 - 2$. If we take $x_1 = 1.5$ (note that $1 < 2^{1/3} < 3$), then Eq. (13.12) gives

$$\begin{aligned} x_2 &= 1.296296296, \\ x_3 &= 1.260932225, \\ x_4 &= 1.259921861, \\ &\left.\begin{aligned} x_5 &= 1.25992105 \\ x_6 &= 1.25992105 \end{aligned}\right\}\text{repetition.} \end{aligned}$$

Thus, to eight-decimal-place accuracy, $2^{1/3} = 1.25992105$.

When applying Newton's method, it is often convenient to replace $f'(x_n)$ by

$$\frac{f(x_n + \delta) - f(x_n)}{\delta},$$

with δ small. Usually $\delta = 0.001$ will give good accuracy. Eq. (13.12) then reads

$$x_{n+1} = x_n - \frac{\delta f(x_n)}{f(x_n + \delta) - f(x_n)}, \quad n = 1, 2, \ldots. \tag{13.13}$$

Example 13.3
Solve the equation $x^2 - 2 = 0$.

Solution: Here $f(x) = x^2 - 2$. Take $x_1 = 1$ and $\delta = 0.001$, then Eq. (13.13) gives

$$x_2 = 1.499750125,$$

$$x_3 = 1.416680519,$$

$$x_4 = 1.414216580,$$

$$x_5 = 1.414213563,$$

$$\left.\begin{array}{l} x_6 = 1.414113562 \\ x_7 = 1.414113562 \end{array}\right\} x_6 = x_7.$$

Numerical integration

Very often definite integrations cannot be done in closed form. When this happens we need some simple and useful techniques for approximating definite integrals. In this section we discuss three such simple and useful methods.

The rectangular rule

The reader is familiar with the interpretation of the definite integral $\int_a^b f(x)dx$ as the area under the curve $y = f(x)$ between the limits $x = a$ and $x = b$:

$$\int_a^b f(x)dx = \sum_{i=1}^{n} f(\alpha_i)(x_i - x_{i-1}),$$

where $x_{i-1} \leq \alpha_i \leq x_i$, $a = x_0 < x_1 < x_2 < \cdots < x_n = b$. We can obtain a good approximation to this definite integral by simply evaluating such an area under the curve $y = f(x)$. We can divide the interval $a \leq x \leq b$ into n subintervals of length $h = (b - a)/n$, and in each subinterval, the function $f(\alpha_i)$ is replaced by a

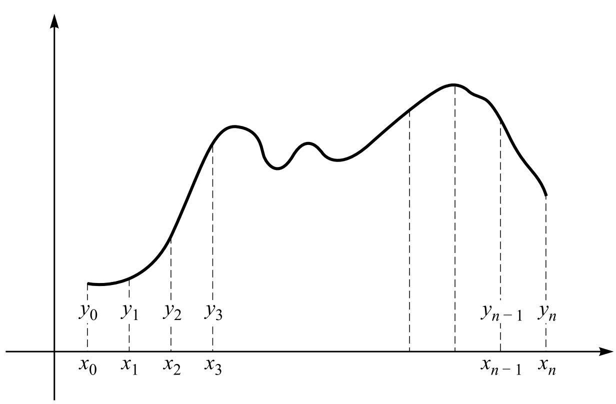

Figure 13.5.

https://doi.org/10.1017/CBO9780511755781.014 Published online by Cambridge University Press

straight line connecting the values at each head or end of the subinterval (or at the center point of the interval), as shown in Fig. 13.5. If we choose the head, $\alpha_i = x_{i-1}$, then we have

$$\int_a^b f(x)dx \approx h(y_0 + y_1 + \cdots + y_{n-1}), \tag{13.14}$$

where $y_0 = f(x_0)$, $y_1 = f(x_1), \ldots, y_{n-1} = f(x_{n-1})$. This method is called the rectangular rule.

It will be shown later that the error decreases as n^2. Thus, as n increases, the error decreases rapidly.

The trapezoidal rule

The trapezoidal rule evaluates the small area of a subinterval slightly differently. The area of a trapezoid as shown in Fig. 13.6 is given by

$$\tfrac{1}{2}h(Y_1 + Y_2).$$

Thus, applied to Fig. 13.5, we have the approximation

$$\int_a^b f(x)dx \approx \frac{(b-a)}{n}\left(\tfrac{1}{2}y_0 + y_1 + y_2 + \cdots + y_{n-1} + \tfrac{1}{2}y_n\right). \tag{13.15}$$

What are the upper and lower limits on the error of this method? Let us first calculate the error for a single subinterval of length $h(= (b-a)/n)$. Writing $x_i + h = z$ and $\varepsilon_i(z)$ for the error, we have

$$\int_{x_i}^{z} f(x)dx = \frac{h}{2}[y_i + y_z] + \varepsilon_i(z),$$

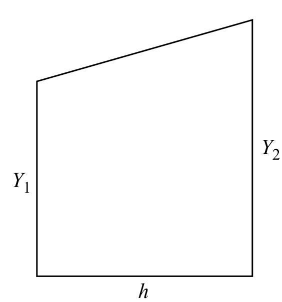

Figure 13.6.

https://doi.org/10.1017/CBO9780511755781.014 Published online by Cambridge University Press

where $y_i = f(x_i)$, $yz = f(z)$. Or

$$\varepsilon_i(z) = \int_{x_i}^{z} f(x)dx - \frac{h}{2}[f(x_i) - f(z)]$$
$$= \int_{x_i}^{z} f(x)dx - \frac{z - x_i}{2}[f(x_i) - f(z)].$$

Differentiating with respect to z:

$$\varepsilon_i'(z) = f(z) - [f(x_i) + f(z)]/2 - (z - x_i)f'(z)/2.$$

Differentiating once again,

$$\varepsilon_i''(z) = -(z - x_i)f''(z)/2.$$

If m_i and M_i are, respectively, the minimum and the maximum values of $f''(z)$ in the subinterval $[x_i, z]$, we can write

$$\frac{z - x_i}{2} m_i \le -\varepsilon_i''(z) \le \frac{z - x_i}{2} M_i.$$

Anti-differentiation gives

$$\frac{(z - x_i)^2}{4} m_i \le -\varepsilon_i'(z) \le \frac{(z - x_i)^2}{4} M_i$$

Anti-differentiation once more gives

$$\frac{(z - x_i)^3}{12} m_i \le -\varepsilon_i(z) \le \frac{(z - x_i)^3}{12} M_i.$$

or, since $z - x_i = h$,

$$\frac{h^3}{12} m_i \le -\varepsilon_i \le \frac{h^3}{12} M_i.$$

If m and M are, respectively, the minimum and the maximum of $f''(z)$ in the interval $[a, b]$ then

$$\frac{h^3}{12} m \le -\varepsilon_i \le \frac{h^3}{12} M \qquad \text{for all } i.$$

Adding the errors for all subintervals, we obtain

$$\frac{h^3}{12} nm \le -\varepsilon \le \frac{h^3}{12} nM$$

or, since $h = (b - a)/n$,

$$\frac{(b - a)^3}{12n^2} m \le -\varepsilon \le \frac{(b - a)^3}{12n^2} M. \tag{13.16}$$

https://doi.org/10.1017/CBO9780511755781.014 Published online by Cambridge University Press

Thus, the error decreases rapidly as n increases, at least for twice-differentiable functions.

Simpson's rule

Simpson's rule provides a more accurate and useful formula for approximating a definite integral. The interval $a \le x \le b$ is subdivided into an even number of subintervals. A parabola is fitted to points a, $a+h$, $a+2h$; another to $a+2h$, $a+3h$, $a+4h$; and so on. The area under a parabola, as shown in Fig. 13.7, is (Problem 13.8)

$$\frac{h}{3}(y_1 + 4y_2 + y_3).$$

Thus, applied to Fig. 13.5, we have the approximation

$$\int_a^b f(x)dx \approx \frac{h}{3}(y_0 + 4y_1 + 2y_2 + 4y_3 + 2y_4 + \cdots + 2y_{n-2} + 4y_{n-1} + y_n), \quad (13.17)$$

with n even and $h = (b-a)/n$.

The analysis of errors for Simpson's rule is fairly involved. It has been shown that the error is proportional to h^4 (or inversely proportional to n^4).

There are other methods of approximating integrals, but they are not so simple as the above three. The method called Gaussian quadrature is very fast but more involved to implement. Many textbooks on numerical analysis cover this method.

Numerical solutions of differential equations

We noted in Chapter 2 that the methods available for the exact solution of differential equations apply only to a few, principally linear, types of differential equations. Many equations which arise in physical science and in engineering are not solvable by such methods and we are therefore forced to find ways of obtaining approximate solutions of these differential equations. The basic idea of approximate solutions is to specify a small increment h and to obtain approximate values of a solution $y = y(x)$ at x_0, $x_0 + h$, $x_0 + 2h$, ...

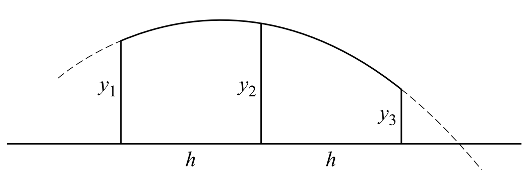

Figure 13.7.

The first-order ordinary differential equation

$$\frac{dy}{dx} = f(x, y), \tag{13.18}$$

with the initial condition $y = y_0$ when $x = x_0$, has the solution

$$y - y_0 = \int_{x_0}^{x} f(t, y(t))dt. \tag{13.19}$$

This integral equation cannot be evaluated because the value of y under the integral sign is unknown. We now consider three simple methods of obtaining approximate solutions: Euler's method, Taylor series method, and the Runge–Kutta method.

Euler's method

Euler proposed the following crude approach to finding the approximate solution. He began at the initial point (x_0, y_0) and extended the solution to the right to the point $x_1 = x_0 + h$, where h is a small quantity. In order to use Eq. (13.19) to obtain the approximation to $y(x_1)$, he had to choose an approximation to f on the interval $[x_0, x_1]$. The simplest of all approximations is to use $f(t, y(t)) = f(x_0, y_0)$. With this choice, Eq. (13.19) gives

$$y(x_1) = y_0 + \int_{x_0}^{x_1} f(x_0, y_0)dt = y_0 + f(x_0, y_0)(x_1 - x_0).$$

Letting $y_1 = y(x_1)$, we have

$$y_1 = y_0 + f(x_0, y_0)(x_1 - x_0). \tag{13.20}$$

From $y_1, y_1' = f(x_1, y_1)$ can be computed. To extend the approximate solution further to the right to the point $x_2 = x_1 + h$, we use the approximation: $f(t, y(t)) = y_1' = f(x_1, y_1)$. Then we obtain

$$y_2 = y(x_2) = y_1 + \int_{x_1}^{x_2} f(x_1, y_1)dt = y_1 + f(x_1, y_1)(x_2 - x_1).$$

Continuing in this way, we approximate y_3, y_4, and so on.

There is a simple geometrical interpretation of Euler's method. We first note that $f(x_0, y_0) = y'(x_0)$, and that the equation of the tangent line at the point (x_0, y_0) to the actual solution curve (or the integral curve) $y = y(x)$ is

$$y - y_0 = \int_{x_0}^{x} f(t, y(t))dt = f(x_0, y_0)(x - x_0).$$

Comparing this with Eq. (13.20), we see that (x_1, y_1) lies on the tangent line to the actual solution curve at (x_0, y_0). Thus, to move from point (x_0, y_0) to point (x_1, y_1) we proceed along this tangent line. Similarly, to move to point (x_2, y_2) we proceed parallel to the tangent line to the solution curve at (x_1, y_1), as shown in Fig. 13.8.

Table 13.1.

x	y (Euler)	y (actual)
1.0	3	3
1.1	3.4	3.43137
1.2	3.861	3.93122
1.3	4.3911	4.50887
1.4	4.99921	5.1745
1.5	5.69513	5.93977
1.6	6.48964	6.81695
1.7	7.39461	7.82002
1.8	8.42307	8.96433
1.9	9.58938	10.2668
2.0	10.9093	11.7463

The merit of Euler's method is its simplicity, but the successive use of the tangent line at the approximate values $y_1, y_2, \ldots$ can accumulate errors. The accuracy of the approximate vale can be quite poor, as shown by the following simple example.

Example 13.4
Use Euler's method to approximate solution to

$$y' = x^2 + y, \quad y(1) = 3 \text{ on interval } [1, 2].$$

Solution: Using $h = 0.1$, we obtain Table 13.1. Note that the use of a smaller step-size h will improve the accuracy.

Euler's method can be improved upon by taking the gradient of the integral curve as the means of obtaining the slopes at x_0 and $x_0 + h$, that is, by using the

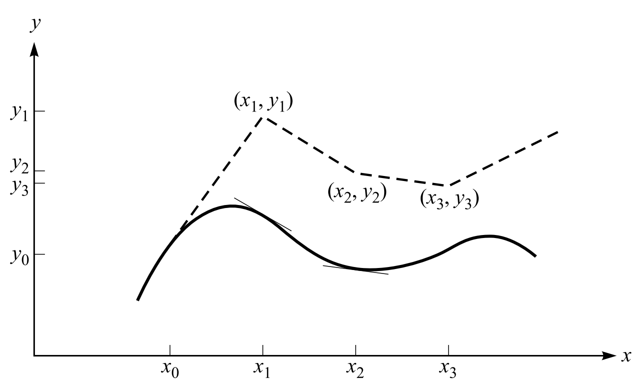

Figure 13.8.

https://doi.org/10.1017/CBO9780511755781.014 Published online by Cambridge University Press

approximate value obtained for y_1, we obtain an improved value, denoted by $(y_1)_1$:

$$(y_1)_1 = y_0 + \tfrac{1}{2}\{f(x_0, y_0) + f(x_0 + h, y_1)\}. \tag{13.21}$$

This process can be repeated until there is agreement to a required degree of accuracy between successive approximations.

The three-term Taylor series method

The rationale for this method lies in the three-term Taylor expansion. Let y be the solution of the first-order ordinary equation (13.18) for the initial condition $y = y_0$ when $x = x_0$ and suppose that it can be expanded as a Taylor series in the neighborhood of x_0. If $y = y_1$ when $x = x_0 + h$, then, for sufficiently small values of h, we have

$$y_1 = y_0 + h\left(\frac{dy}{dx}\right)_0 + \frac{h^2}{2!}\left(\frac{d^2y}{dx^2}\right)_0 + \frac{h^3}{3!}\left(\frac{d^3y}{dx^3}\right)_0 + \cdots. \tag{13.22}$$

Now

$$\frac{dy}{dx} = f(x, y),$$

$$\frac{d^2y}{dx^2} = \frac{\partial f}{\partial x} + \frac{dy}{dx}\frac{\partial f}{\partial y} = \frac{\partial f}{\partial x} + f\frac{\partial f}{\partial y},$$

and

$$\begin{aligned}\frac{d^3y}{dx^3} &= \left(\frac{\partial}{\partial x} + f\frac{\partial}{\partial y}\right)\left(\frac{\partial f}{\partial x} + f\frac{\partial f}{\partial y}\right)\\ &= \frac{\partial^2 f}{\partial x^2} + \frac{\partial f}{\partial y}\frac{\partial f}{\partial y} + 2f\frac{\partial^2 f}{\partial x \partial y} + f\left(\frac{\partial f}{\partial y}\right)^2 + f^2\frac{\partial^2 f}{\partial y^2}.\end{aligned}$$

Equation (13.22) can be rewritten as

$$y_1 = y_0 + hf(x_0, y_0) + \frac{h^2}{2}\left[\frac{\partial f(x_0, y_0)}{\partial x} + f(x_0, y_0)\frac{\partial f(x_0, y_0)}{\partial y}\right],$$

where we have dropped the h^3 term. We now use this equation as an iterative equation:

$$y_{n+1} = y_n + hf(x_n, y_n) + \frac{h^2}{2}\left[\frac{\partial f(x_n, y_n)}{\partial x} + f(x_n, y_n)\frac{\partial f(x_n, y_n)}{\partial y}\right]. \tag{13.23}$$

That is, we compute $y_1 = y(x_0 + h)$ from y_0, $y_2 = y(x_1 + h)$ from y_1 by replacing x by x_1, and so on. The error in this method is proportional h^3. A good approxima-

https://doi.org/10.1017/CBO9780511755781.014 Published online by Cambridge University Press

Table 13.2.

n	x_n	y_n	y_{n+1}
0	1.0	−2.0	−2.1
1	1.1	−2.1	−2.2
2	1.2	−2.2	−2.3
3	1.3	−2.3	−2.4
4	1.4	−2.4	−2.5
5	1.5	−2.5	−2.6

tion can be obtained for y_n by summing a number of terms of the Taylor's expansion. To illustrate this method, let us consider a very simple example.

Example 13.5
Find the approximate values of y_1 through y_{10} for the differential equation $y' = x + y$, with the initial condition $x_0 = 1.0$ and $y = -2.0$.

Solution: Now $f(x, y) = x + y$, $\partial f/\partial x = \partial f/\partial y = 1$ and Eq. (13.23) reduces to

$$y_{n+1} = y_n + h(x_n + y_n) + \frac{h^2}{2}(1 + x_n + y_n).$$

Using this simple formula with $h = 0.1$ we obtain the results shown in Table 13.2.

The Runge–Kutta method

In practice, the Taylor series converges slowly and the accuracy involved is not very high. Thus we often resort to other methods of solution such as the Runge–Kutta method, which replaces the Taylor series, Eq. (13.23), with the following formula:

$$y_{n+1} = y_n + \frac{h}{6}(k_1 + 4k_2 + k_3), \tag{13.24}$$

where

$$k_1 = f(x_n, y_n), \tag{13.24a}$$

$$k_2 = f(x_n + h/2,\ y_n + hk_1/2), \tag{13.24b}$$

$$k_3 = f(x_n + h, y_0 + 2hk_2 - hk_1). \tag{13.24c}$$

This approximation is equivalent to Simpson's rule for the approximate integration of $f(x, y)$, and it has an error proportional to h^4. A beauty of the

https://doi.org/10.1017/CBO9780511755781.014 Published online by Cambridge University Press

Runge–Kutta method is that we do not need to compute partial derivatives, but it becomes rather complicated if pursued for more than two or three steps.

The accuracy of the Runge–Kutta method can be improved with the following formula:

$$y_{n+1} = y_n + \frac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4), \tag{13.25}$$

where

$$k_1 = f(x_n, y_n), \tag{13.25a}$$

$$k_2 = f(x_n + h/2, y_n + hk_1/2), \tag{13.25b}$$

$$k_3 = f(x_n + h, y_0 + hk_2/2), \tag{13.25c}$$

$$k_4 = f(x_n + h, y_n + hk_3). \tag{13.25d}$$

With this formula the error in y_{n+1} is of order h^5.

You may wonder how these formulas are established. To this end, let us go back to Eq. (13.22), the three-term Taylor series, and rewrite it in the form

$$\begin{aligned} y_1 = y_0 + hf_0 + (1/2)h^2(A_0 + f_0B_0) + (1/6)h^3(C_0 + 2f_0D_0 + f_0^2E_0 \\ + A_0B_0 + f_0B_0^2) + O(h^4), \end{aligned} \tag{13.26}$$

where

$$A = \frac{\partial f}{\partial x}, \quad B = \frac{\partial f}{\partial y}, \quad C = \frac{\partial^2 f}{\partial x^2}, \quad D = \frac{\partial^2 f}{\partial x \partial y}, \quad E = \frac{\partial^2 f}{\partial y^2}$$

and the subscript 0 denotes the values of these quantities at (x_0, y_0).

Now let us expand k_1, k_2, and k_3 in the Runge–Kutta formula (13.24) in powers of h in a similar manner:

$$k_1 = hf(x_0, y_0),$$

$$\begin{aligned} k_2 &= f(x_0 + h/2, y_0 + k_1h/2), \\ &= f_0 + \frac{1}{2}h(A_0 + f_0B_0) + \frac{1}{8}h^2(C_0 + 2f_0D_0 + f_0^2E_0) + O(h^3). \end{aligned}$$

Thus

$$2k_2 - k_1 = f_0 + h(A_0 + f_0B_0) + \cdots$$

and

$$\left(\frac{d}{dh}(2k_2 - k_1)\right)_{h=0} = f_0, \left(\frac{d^2}{dh^2}(2k_2 - k_1)\right)_{h=0} = 2(A_0 + f_0B_0).$$

Then

$$\begin{aligned}k_3 &= f(x_0 + h, y_0 + 2hk_2 - hk_1)\\ &= f_0 + h(A_0 + f_0 B_0) + (1/2)h^2\{C_0 + 2f_0 D_0 + f_0^2 E_0 + 2B_0(A_0 + f_0 B_0)\}\\ &\quad + O(h^3).\end{aligned}$$

and

$$\begin{aligned}(1/6)(k_1 + 4k_2 + k_3) &= hf_0 + (1/2)h^2(A_0 + f_0 B_0)\\ &\quad + (1/6)h^3(C_0 + 2f_0 D_0 + f_0^2 E_0 + A_0 B_0 + f_0 B_0^2) + O(h^4).\end{aligned}$$

Comparing this with Eq. (13.26), we see that it agrees with the Taylor series expansion (up to the term in h^3) and the formula is established. Formula (13.25) can be established in a similar manner by taking one more term of the Taylor series.

Example 13.6
Using the Runge–Kutta method and $h = 0.1$, solve

$$y' = x - y^2/10; \quad x_0 = 0, \quad y_0 = 1.$$

Solution: With $h = 0.1$, $h^4 = 0.0001$ and we may use the Runge–Kutta third-order approximation.

First step: $x_0 = 0,\ y_0 = 0,\ f_0 = -0.1,$
$k_1 = -0.1,\ y_0 + hk_1/2 = 0.995,$
$k_2 = -0.049,\ 2k_2 - k_1 = 0.002,\ k_3 = 0,$
$$y_1 = y_0 + \frac{h}{6}(k_1 + 4k_2 + k_1) = 0.9951.$$

Second step: $x_1 = x_0 + h = 0.1,\ y_1 = 0.9951,\ f_1 = 0.001,$
$k_1 = 0.001,\ y_1 + hk_1/2 = 0.9952,$
$k_2 = 0.051,\ 2k_2 - k_1 = 0.101,\ k_3 = 0.099,$
$$y_2 = y_1 + \frac{h}{6}(k_1 + 4k_2 + k_1) = 1.0002.$$

Third step: $x_2 = x_1 + h = 0.2,\ y_2 = 1.0002,\ f_2 = 0.1,$
$k_1 = 0.1,\ y_2 + hk_1/2 = 1.0052,$
$k_2 = 0.149,\ 2k_2 - k_1 = 0.198,\ k_3 = 0.196,$
$$y_3 = y_2 + \frac{h}{6}(k_1 + 4k_2 + k_1) = 1.0151.$$

Equations of higher order. System of equations

The methods in the previous sections can be extended to obtain numerical solutions of equations of higher order. An nth-order differential equation is equivalent to n first-order differential equations in $n+1$ variables. Thus, for instance, the second-order equation

$$y'' = f(x, y, y'), \tag{13.27}$$

with initial conditions

$$y(x_0) = y_0, \quad y'(x_0) = y_0', \tag{13.28}$$

can be written as a system of two equations of first order by setting

$$y' = u, \tag{13.29}$$

then Eqs. (13.27) and (13.28) become

$$u' = f(x, y, u), \tag{13.30}$$

$$y(x_0) = y_0, \quad u(x_0) = u_0. \tag{13.31}$$

The two first-order equations (13.29) and (13.30) with the initial conditions (13.31) are completely equivalent to the original second-order equation (13.27) with the initial conditions (13.28). And the methods in the previous sections for determining approximate solutions can be extended to solve this system of two first-order equations. For example, the equation

$$y'' - y = 2,$$

with initial conditions

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

is equivalent to the system

$$y' = x + u, \quad u' = 1 + y,$$

with

$$y(0) = -1, \quad u(0) = 1.$$

These two first-order equations can be solved with Taylor's method (Problem 13.12).

The simple methods outlined above all have the disadvantage that the error in approximating to values of y is to a certain extent cumulative and may become large unless some form of checking process is included. For this reason, methods of solution involving finite difference are devised, most of them being variations of the Adams–Bashforth method that contains a self-checking process. This method

https://doi.org/10.1017/CBO9780511755781.014 Published online by Cambridge University Press

is quite involved and because of limited space we shall not cover it here, but it is discussed in any standard textbook on numerical analysis.

Least-squares fit

We now look at the problem of fitting of experimental data. In some experimental situations there may be underlying theory that suggests the kind of function to be used in fitting the data. Often there may be no theory on which to rely in selecting a function to represent the data. In such circumstances a polynomial is often used. We saw earlier that the $m+1$ coefficients in the polynomial

$$y = a_0 + a_1 x + \cdots + a_m x^m$$

can always be determined so that a given set of $m+1$ points (x_i, y_i), where the xs may be unequal, lies on the curve described by the polynomial. However, when the number of points is large, the degree m of the polynomial is high, and an attempt to fit the data by using a polynomial is very laborious. Furthermore, the experimental data may contain experimental errors, and so it may be more sensible to represent the data approximately by some function $y = f(x)$ that contains a few unknown parameters. These parameters can then be determined so that the curve $y = f(x)$ fits the data. How do we determine these unknown parameters?

Let us represent a set of experimental data (x_i, y_i), where $i = 1, 2, \ldots, n$, by some function $y = f(x)$ that contains r parameters $a_1, a_2, \ldots, a_r$. We then take the deviations (or residuals)

$$d_i = f(x_i) - y_i \tag{13.32}$$

and form the weighted sum of squares of the deviations

$$S = \sum_{i=1}^{n} w_i (d_i)^2 = \sum_{i=1}^{n} w_i [f(x_i) - y_i]^2, \tag{13.33}$$

where the weights w_i express our confidence in the accuracy of the experimental data. If the points are equally weighted, the ws can all be set to 1.

It is clear that the quantity S is a function of as: $S = S(a_1, a_2, \ldots, a_r)$. We can now determine these parameters so that S is a minimum:

$$\frac{\partial S}{\partial a_1} = 0, \quad \frac{\partial S}{\partial a_2} = 0, \ \ldots, \ \frac{\partial S}{\partial a_r} = 0. \tag{13.34}$$

The set of r equations (13.34) is called the normal equations and serves to determine the r unknown as in $y = f(x)$. This particular method of determining the unknown as is known as the method of least squares.

We now illustrate the construction of the normal equations with the simplest case: $y = f(x)$ is a linear function:

$$y = a_1 + a_2 x. \tag{13.35}$$

The deviations d_i are given by

$$d_i = (a_1 + a_2 x) - y_i$$

and so, assuming $w_i = 1$

$$S = \sum_{i=1}^{n} d_i^2 = (a_1 + a_2 x_1 - y_1)^2 + (a_1 + a_2 x_2 - y_2)^2 + \cdots + (a_1 + a_2 x_r - y_r)^2.$$

We now find the partial derivatives of S with respect to a_1 and a_2 and set these to zero:

$$\partial S/\partial a_1 = 2(a_1 + a_2 x_1 - y_1) + 2(a_1 + a_2 x_2 - y_2) + \cdots + 2(a_1 + a_2 x_n - y_n) = 0,$$

$$\partial S/\partial a_2 = 2x_1(a_1 + a_2 x_1 - y_1) + 2x_2(a_1 + a_2 x_2 - y_2) + \cdots + 2x_n(a_1 + a_2 x_n - y_n) = 0.$$

Dividing out the factor 2 and collecting the coefficients of a_1 and a_2, we obtain

$$n a_1 + \left(\sum_{i=1}^{n} x_i\right) a_2 = \sum_{i=1}^{n} y_1, \tag{13.36}$$

$$\left(\sum_{i=1}^{n} x_i\right) a_1 + \left(\sum_{i=1}^{n} x_i^2\right) a_2 = \sum_{i=1}^{n} x_i y_i. \tag{13.37}$$

These equations can be solved for a_1 and a_2.

Problems

13.1. Given six points $(-1, 0)$, $(-0.8, 2)$, $(-0.6, 1)$, $(-0.4, -1)$, $(-0.2, 0)$, and $(0, -4)$, determine a smooth function $y = f(x)$ such that $y_i = f(x_i)$.

13.2. Find an approximate value of the real root of

$$x - \tan x = 0$$

near $x = 3\pi/2$.

13.3. Find the angle subtended at the center of a circle by an arc whose length is double the length of the chord.

13.4. Use Newton's method to solve

$$e^{x^2} - x^3 + 3x - 4 = 0,$$

with $x_0 = 0$ and $h = 0.001$.

13.5. Use Newton's method to find a solution of

$$\sin(x^3 + 2) = 1/x,$$

with $x_0 = 1$ and $h = 0.001$.

13.6. Approximate the following integrals using the rectangular rule, the trapezoidal rule, and Simpson's rule, with $n = 2, 4, 10, 20, 50$:

(*a*) $\displaystyle\int_0^{\pi/2} e^{-x^2} \sin(x^2 + 1)dx;$

(*b*) $\displaystyle\int_0^{\sqrt{2}} \frac{\sin(x^2) + 3x - 2}{x + 4} dx;$

(*c*) $\displaystyle\int_0^1 \frac{dx}{\sqrt{2 - \sin^2 x}}.$

13.7 Show that the area under a parabola, as shown in Fig. 13.7, is given by

$$A = \frac{h}{3}(y_1 + 4y_2 + y_3).$$

13.8. Using the improved Euler's method, find the value of y when $x = 0.2$ on the integral curve of the equation $y' = x^2 - 2y$ through the point $x = 0$, $y = 1$.

13.9. Using Taylor's method, find correct to four places of decimals values of y corresponding to $x = 0.2$ and $x = -0.2$ for the solution of the differential equation

$$dy/dx = x - y^2/10,$$

with the initial condition $y = 1$ when $x = 0$.

13.10. Using the Runge–Kutta method and $h = 0.1$, solve

$$y' = x^2 - \sin(y^2), \quad x_0 = 1 \text{ and } y_0 = 4.7.$$

13.11. Using the Runge–Kutta method and $h = 0.1$, solve

$$y' = ye^{-x^2}, \quad x_0 = 1 \text{ and } y_0 = 3.$$

13.12. Using Taylor's method, obtain the solution of the system

$$y' = x + u, \quad u' = 1 + y$$

with

$$y(0) = -1, \quad u(0) = 1.$$

.

13.13. Find to four places of decimals the solution between $x = 0$ and $x = 0.5$ of the equations

$$y' = \tfrac{1}{2}(y + u), \quad u' = \tfrac{1}{2}(y^2 - u^2),$$

with $y = u = 1$ when $x = 0$.

https://doi.org/10.1017/CBO9780511755781.014 Published online by Cambridge University Press

13.14. Find to three places of decimals a solution of the equation

$$y'' + 2xy' - 4y = 0,$$

with $y = y' = 1$ when $x = 0$.

13.15. Use Eqs. (13.36) and (13.37) to calculate the coefficients in $y = a_1 + a_2x$ to fit the following data: $(x, y) = (1, 1.7), (2, 1.8), (3, 2.3), (4, 3.2)$.

https://doi.org/10.1017/CBO9780511755781.014 Published online by Cambridge University Press