5 Numerical Integration

In electrical field theory, it is proved that the magnetic field induced by a current flowing in a circular loop of wire has intensity

$$H(x) = \frac{4Ir}{r^2 - x^2}\int_0^{\pi/2}\left[1 - \left(\frac{x}{r}\right)^2 \sin^2\theta\right]^{1/2} d\theta$$

where I is the current, r is the radius of the loop, and x is the distance from the center to the point where the magnetic intensity is being computed $(0 \leqq x \leqq r)$. If I, r, and x are given, we have a formidable integral to evaluate. It is an **elliptic integral** and not expressible in terms of familiar functions. But H *can* be computed precisely by the methods of this chapter. For example, if $I = 15.3$, $r = 120$, and $x = 84$, we find $H = 1.35566\,1135$ accurate to nine decimals.

5.1 Lower and Upper Sums

Elementary calculus focuses largely on two important processes of mathematics: differentiation and integration. In Section 1.1, numerical differentiation was considered briefly; it was taken up again in Section 4.3. In this chapter, the process of integration is examined from the standpoint of numerical mathematics.

Definite and Indefinite Integrals

It is customary to distinguish two types of integrals: the definite and the indefinite integral. The **indefinite integral** of a function is another *function* or a class of functions, whereas the **definite integral** of a function over a fixed interval is a *number*. For example,

$$\text{Indefinite integral:} \quad \int x^2\,dx = \frac{1}{3}x^3 + C$$

$$\text{Definite integral:} \quad \int_0^2 x^2\,dx = \frac{8}{3}$$

Actually, a function has not just one but many indefinite integrals. These differ from each other by constants. Thus, in the preceding example, any constant value may be assigned to C, and the result is still an indefinite integral. In elementary calculus, the concept of an indefinite integral is identical with the concept of an antiderivative. An **antiderivative** of a function f is any function F having the property that $F' = f$.

The definite and indefinite integrals are related by the **Fundamental Theorem of Calculus,*** which states that $\int_a^b f(x)\,dx$ can be computed by first finding an antiderivative F of f and then evaluating $F(b) - F(a)$. Thus, using traditional notation, we have

$$\int_1^3 (x^2 - 2)\,dx = \left(\frac{x^3}{3} - 2x\right)\Bigg|_1^3 = \left(\frac{27}{3} - 6\right) - \left(\frac{1}{3} - 2\right) = \frac{14}{3}$$

As another example of the Fundamental Theorem of Calculus, we can write

$$\int_a^b F'(x)\,dx = F(b) - F(a)$$

$$\int_a^x F'(t)\,dt = F(x) - F(a)$$

If this second equation is differentiated with respect to x, the result is (and here we have put $f = F'$)

$$\frac{d}{dx}\int_a^x f(t)\,dt = f(x)$$

This last equation shows that $\int_a^x f(t)\,dt$ must be an antiderivative (indefinite integral) of f.

The foregoing technique for computing definite integrals is virtually the only one emphasized in elementary calculus. The definite integral of a function, however, has an interpretation as the area under a curve, and so the existence of a numerical value for $\int_a^b f(x)\,dx$ should not depend logically on our limited ability to find antiderivatives. Thus, for instance,

$$\int_0^1 e^{x^2}\,dx$$

has a precise numerical value despite the fact that there is no elementary function F such that $F'(x) = e^{x^2}$. By the preceding remarks, e^{x^2} does have antiderivatives, one of which is

$$F(x) = \int_0^x e^{t^2}\,dt$$

However, this form of the function F is of no help in determining the numerical value sought.

Lower and Upper Sums

The existence of the definite integral of a nonnegative function f on a closed interval $[a, b]$ is based on an interpretation of that integral as the area under the graph of f. The definite integral is defined by means of two concepts, the *lower sums* of f and the *upper sums* of f; these are approximations to the area under the graph.

***Fundamental Theorem of Calculus**: If f is continuous on the interval $[a, b]$ and F is an antiderivative of f, then

$$\int_a^b f(x)\,dx = F(b) - F(a)$$

Let P be a **partition** of the interval $[a, b]$ given by

$$P = \{a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b\}$$

with partition points $x_0, x_1, x_2, \ldots, x_n$ that divide the interval $[a, b]$ into n subintervals $[x_i, x_{i+1}]$. Now denote by m_i the **greatest lower bound** (*infimum* or inf) of $f(x)$ on the subinterval $[x_i, x_{i+1}]$. In symbols,

$$m_i = \inf\{f(x) : x_i \leqq x \leqq x_{i+1}\}$$

Likewise, we denote by M_i the **least upper bound** (*supremum* or sup) of $f(x)$ on $[x_i, x_{i+1}]$. Thus,

$$M_i = \sup\{f(x) : x_i \leqq x \leqq x_{i+1}\}$$

The **lower sums** and **upper sums** of f corresponding to the given partition P are defined to be

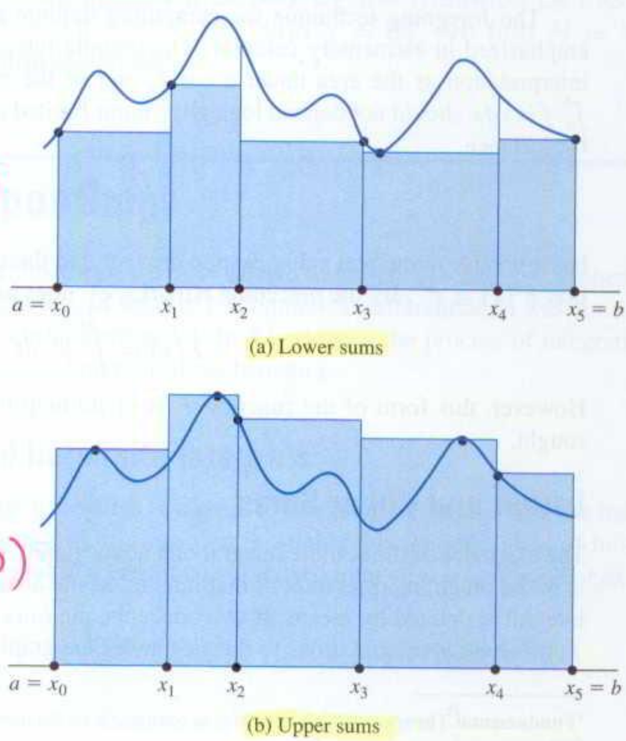
$$L(f; P) = \sum_{i=0}^{n-1} m_i(x_{i+1} - x_i)$$

$$U(f; P) = \sum_{i=0}^{n-1} M_i(x_{i+1} - x_i)$$

If f is a positive function, these two quantities can be interpreted as estimates of the area under the curve for f. These sums are shown in Figure 5.1.

L(f;P)

U(f;P)



(a) Lower sums

(b) Upper sums

FIGURE 5.1 Illustrating lower and upper sums

EXAMPLE 1 What are the numerical values of the upper and lower sums for $f(x) = x^2$ on the interval $[0, 1]$ if the partition is $P = \left\{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\right\}$?

Solution We want the value of

$$U(f;P) = M_0(x_1 - x_0) + M_1(x_2 - x_1) + M_2(x_3 - x_2) + M_3(x_4 - x_3)$$

Since f is increasing on $[0, 1]$, $M_0 = f(x_1) = \frac{1}{16}$. Similarly, $M_1 = f(x_2) = \frac{1}{4}$, $M_2 = f(x_3) = \frac{9}{16}$, and $M_3 = f(x_4) = 1$. The widths of the subintervals are all equal to $\frac{1}{4}$. Hence,

$$U(f;P) = \tfrac{1}{4}\left(\tfrac{1}{16} + \tfrac{1}{4} + \tfrac{9}{16} + 1\right) = \tfrac{15}{32}$$

In the same way, we find that $m_0 = f(x_0) = 0$, $m_1 = \frac{1}{16}$, $m_2 = \frac{1}{4}$, and $m_3 = \frac{9}{16}$. Hence,

$$L(f;P) = \tfrac{1}{4}\left(0 + \tfrac{1}{16} + \tfrac{1}{4} + \tfrac{9}{16}\right) = \tfrac{7}{32}$$

If we had no other way of calculating $\int_0^1 x^2\,dx$, we would take a value halfway between $U(f;P)$ and $L(f;P)$ as the best estimate. This number is $\frac{11}{32}$. The correct value is $\frac{1}{3}$, and the error is $\frac{11}{32} - \frac{1}{3} = \frac{1}{96}$. ■

It is intuitively clear that the upper sum *overestimates* the area under the curve, and the lower sum *underestimates* it. Therefore, the expression $\int_a^b f(x)\,dx$, which we are trying to define, is *required* to satisfy the basic inequality

$$L(f;P) \leqq \int_a^b f(x)\,dx \leqq U(f;P) \tag{1}$$

for all partitions P. It turns out that if f is a *continuous* function defined on $[a, b]$, then Inequality (1) does indeed define the integral. That is, there is one and only one real number that is greater than or equal to all lower sums of f and less than or equal to all upper sums of f. This unique number (depending on f, a, and b) is defined to be $\int_a^b f(x)\,dx$. The integral also exists if f is monotone increasing on $[a, b]$ or monotone decreasing on $[a, b]$.

Riemann-Integrable Functions

We consider the least upper bound (supremum) of the set of all numbers $L(f;P)$ obtained when P is allowed to range over all partitions of the interval $[a, b]$. This is abbreviated $\sup_P L(f;P)$. Similarly, we consider the greatest lower bound (infimum) of $U(f;P)$ when P ranges over all partitions of $[a, b]$. This is denoted by $\inf_P U(f;P)$. Now if these two numbers are the same—that is, if

$$\inf_P U(f;P) = \sup_P L(f;P) \tag{2}$$

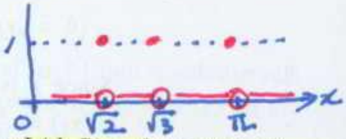
then we say that f is **Riemann-integrable** on $[a, b]$ and define $\int_a^b f(x)\,dx$ to be the common value obtained in Equation (2). The important result mentioned above can be stated formally as follows:

■ THEOREM 1 **THEOREM ON RIEMANN INTEGRAL**

Every continuous function defined on a closed and bounded interval of the real line is Riemann-integrable.

There are plenty of functions that are *not* Riemann-integrable. The simplest is known as the **Dirichlet function**:

$$d(x) = \begin{cases} 0 & \text{if } x \text{ is rational} \\ 1 & \text{if } x \text{ is irrational} \end{cases}$$



For any interval $[a, b]$ and for any partition P of $[a, b]$, we have $L(d; P) = 0$ and $U(d; P) = b - a$. Hence,

$$0 = \sup_P L(d; P) < \inf_P U(d; P) = b - a$$

In calculus, it is proved not only that the Riemann integral of a continuous function on $[a, b]$ exists but also that it can be obtained by two limits:

$$\lim_{n\to\infty} L(f; P_n) = \int_a^b f(x)\,dx = \lim_{n\to\infty} U(f; P_n)$$

in which $P_0, P_1, \ldots$ is any sequence of partitions with the property that the length of the largest subinterval in P_n converges to zero as $n \to \infty$. Furthermore, if it is so arranged that P_{n+1} is obtained from P_n by adding new points (and not deleting points), then the lower sums converge *upward* to the integral and the upper sums converge *downward* to the integral. From the numerical standpoint, this is a desirable feature of the process because at each step, an interval that contains the unknown number $\int_a^b f(x)\,dx$ will be available. Moreover, these intervals shrink in width at each succeeding step.

Examples and Pseudocode

The process just described can easily be carried out on a computer. To illustrate, we select the function $f(x) = e^{-x^2}$ and the interval $[0, 1]$; that is, we consider

$$\int_0^1 e^{-x^2}\,dx \tag{3}$$

This function is of great importance in statistics, but its indefinite integral cannot be obtained by the elementary techniques of calculus. For partitions, we take equally spaced points in $[0, 1]$. Thus, if there are to be n subintervals in P_n, then we define $P_n = \{x_0, x_1, \ldots, x_n\}$, where $x_i = ih$ for $0 \leqq i \leqq n$ and $h = 1/n$. Since e^{-x^2} is *decreasing* on $[0, 1]$, the least value of f on the subinterval $[x_i, x_{i+1}]$ occurs at x_{i+1}. Similarly, the greatest value occurs at x_i. Hence, $m_i = f(x_{i+1})$ and $M_i = f(x_i)$. Putting this into the formulas for the upper and lower sums, we obtain for this function

$$L(f; P_n) = \sum_{i=0}^{n-1} hf(x_{i+1}) = h\sum_{i=0}^{n-1} e^{-x_{i+1}^2}$$

$$U(f; P_n) = \sum_{i=0}^{n-1} hf(x_i) = h\sum_{i=0}^{n-1} e^{-x_i^2}$$

Since these sums are almost the same, it is more economical to compute $L(f; P_n)$ by the given formula and to obtain $U(f; P_n)$ by observing that

$$U(f; P_n) = hf(x_0) + L(f; P_n) - hf(x_n) = L(f; P_n) + h(1 - e^{-1})$$

For historical reasons, formulas for approximating definite integrals are called **rules**. The lower and upper sums give rise to the left and right rectangle rules, the midpoint rule, the trapezoid rule, and many other rules, some of which are found in the problems and subsequent chapters of this book. A large collection of these **quadrature rules** can be found in Abramowitz and Stegun [1964], *Standard Mathematical Tables*, which had its origins in a U.S. government work project conducted during the Depression of the 1930s.

The word **quadrature** has several meanings both in mathematics and in astronomy. In the dictionary, the first mathematical meaning is the process of finding a square whose area is equal to the area enclosed by a given curve. The general mathematical meaning is the process of determining the area of a surface, especially one bounded by a curve. We use it primarily to mean the approximation of the area under a curve using a numerical integration procedure.

Summary

(1) Let $P = \{a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b\}$ be a **partition** of the interval $[a, b]$, which divides the interval $[a, b]$ into n subintervals $[x_i, x_{i+1}]$. The **lower sums** and **upper sums** of f corresponding to the given partition P are

$$L(f; P) = \sum_{i=0}^{n-1} m_i(x_{i+1} - x_i)$$
$$U(f; P) = \sum_{i=0}^{n-1} M_i(x_{i+1} - x_i)$$

where m_i is the **greatest lower bound** and M_i is the **least upper bound** of $f(x)$ on the subinterval $[x_i, x_{i+1}]$, namely,

$$m_i = \inf\{f(x) : x_i \leqq x \leqq x_{i+1}\}$$
$$M_i = \sup\{f(x) : x_i \leqq x \leqq x_{i+1}\}$$

(2) We have

$$L(f; P) \leqq \int_a^b f(x)\,dx \leqq U(f; P)$$

Problems 5.1

[a]**1.** If we estimate $\int_0^1 (x^2+2)^{-1}\,dx$ by means of a lower sum using the partition $P = \{0, \frac{1}{2}, 1\}$, what is the result?

2. What is the result if we estimate $\int_1^2 x^{-1}\,dx$ by means of the upper sum using the partition $P = \{1, \frac{3}{2}, 2\}$?

[a]**3.** Calculate an approximate value of $\int_0^\alpha [(e^x - 1)/x]\,dx$ for $\alpha = 10^{-4}$ correct to 14 decimal places (rounded). *Hint:* Use Taylor series.

5.2 Trapezoid Rule

The next method considered is an improvement over the coarse method of the preceding section. Moreover, it is an important ingredient of the Romberg algorithm of the next section.

This method is called the **trapezoid rule** and is based on an estimation of the area beneath a curve using trapezoids. Again, the estimation of $\int_a^b f(x)\,dx$ is approached by first dividing the interval $[a, b]$ into subintervals according to the *partition* $P = \{a = x_0 < x_1 < x_2 < \cdots < x_n = b\}$. For each such partition of the interval (the partition points x_i need not be uniformly spaced), an estimation of the integral by the trapezoid rule is obtained. We denote it by $T(f; P)$. Figure 5.2 shows what the trapezoids are. A typical trapezoid has the subinterval $[x_i, x_{i+1}]$ as its base, and the two vertical sides are $f(x_i)$ and $f(x_{i+1})$ (see Figure 5.3). The area is equal to the base times the average height, and we have the **basic trapezoid rule** for the subinterval $[x_i, x_{i+1}]$:

$$\int_{x_i}^{x_{i+1}} f(x)\,dx \approx \frac{1}{2}(x_{i+1} - x_i)[f(x_i) + f(x_{i+1})]$$

Hence, the total area of all the trapezoids is

$$\int_a^b f(x)\,dx \approx \frac{1}{2}\sum_{i=0}^{n-1}(x_{i+1} - x_i)[f(x_i) + f(x_{i+1})]$$

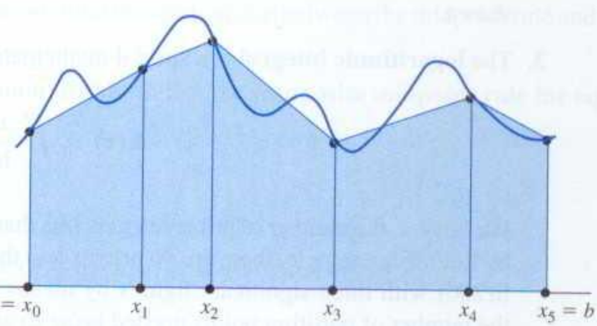


FIGURE 5.2 Trapezoid rule

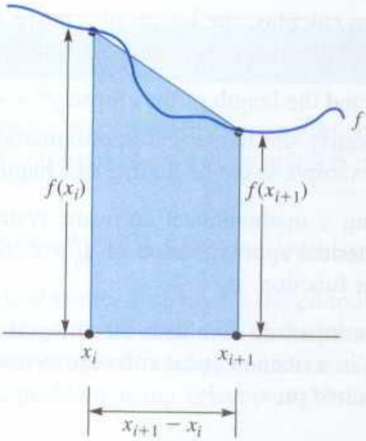


FIGURE 5.3 Typical trapezoid

This formula is called the **composite trapezoid rule**. The composite trapezoid rule is easy to understand: on each subinterval $[x_i, x_{i+1}]$, we multiply $(x_{i+1} - x_i)$ times the average of $f(x_i)$ and $f(x_{i+1})$.

Uniform Spacing

In practice and in the Romberg algorithm (discussed in the next section), the trapezoid rule is used with a *uniform* partition of the interval. This means that the division points x_i are equally spaced: $x_i = a + ih$, where $h = (b - a)/n$ and $0 \leqq i \leqq n$. Think of h as the step size in the process. In this case, the formula for $T(f; P)$ can be given in simpler form because $x_{i+1} - x_i = h$. Thus, we obtain

$$T(f; P) = \frac{h}{2} \sum_{i=0}^{n-1} [f(x_i) + f(x_{i+1})]$$

It should be emphasized that to economize the amount of arithmetic, the computationally preferable formula for the composite trapezoid rule is

$$\int_a^b f(x)\,dx \approx T(f; P) = h\left\{\frac{1}{2}[f(x_0) + f(x_n)] + \sum_{i=1}^{n-1} f(x_i)\right\} \tag{1}$$

Here, we have expanded the summation and gathered similar terms in the new summation. To illustrate, we consider the integral

$$\int_0^1 e^{-x^2}\,dx$$

which was approximated in Section 5.1 by lower and upper sums. Here is a pseudocode for Equation (1) with $n = 60$ and $f(x) = e^{-x^2}$:

```
program Trapezoid
integer i;   real h, sum, x
integer n ← 60;   real a ← 0, b ← 1
h ← (b − a)/n
sum ← ½[f(a) + f(b)]
for i = 1 to n − 1 do
    x ← a + ih
    sum ← sum + f(x)
end for
sum ← (sum)h
output sum
end Trapezoid

real function f(x)
real x
f ← 1/e^(x²)
end function f
```

The computer output for the approximate value of the integral is 0.74681.

EXAMPLE 1 Compute

$$\int_0^1 (\sin x/x)\,dx$$

by using the composite trapezoid rule with six uniform points (cf. Computer Problem 5.1.2).

Solution The function values are arranged in a table as follows:

x_i	$f(x_i)$
0.0	1.00000
0.2	0.99335
0.4	0.97355
0.6	0.94107
0.8	0.89670
1.0	0.84147

Notice that we have assigned the value $\sin x/x = 1$ at $x = 0$. Then

$$\begin{aligned} T(f;P) &= 0.2\sum_{i=1}^{4} f(x_i) + (0.1)[f(x_0) + f(x_5)] \\ &= (0.2)(3.80467) + (0.1)(1.84147) \\ &= 0.94508 \end{aligned}$$

This result is not accurate to all the digits shown, as might be expected because only five subintervals were used. Using mathematical software, we obtain $\mathrm{Si}(1) \approx 0.94608\,30704$. (Refer to Computer Problem 5.1.2.) We shall see later how to determine a suitable value for n to obtain a desired accuracy using the trapezoid rule. ■

Error Analysis

The next task is to analyze the error incurred in using the trapezoid rule to estimate an integral. We shall establish the following result.

■ THEOREM 1 **THEOREM ON PRECISION OF TRAPEZOID RULE**

If f'' exists and is continuous on the interval $[a, b]$ and if the composite trapezoid rule T with uniform spacing h is used to estimate the integral $I = \int_a^b f(x)\,dx$, then for some ζ in (a, b),

$$I - T = -\tfrac{1}{12}(b-a)h^2 f''(\zeta) = \mathcal{O}(h^2)$$

Proof The first step in the analysis is to prove the above result when $a = 0$, $b = 1$, and $h = 1$. In this case, we have to show that

$$\int_0^1 f(x)\,dx - \frac{1}{2}[f(0) + f(1)] = -\frac{1}{12}f''(\zeta) \tag{2}$$

This is easily established with the aid of the error formula for polynomial interpolation (see Section 4.2). To use this formula, let p be the polynomial of degree 1 that interpolates f at

Multidimensional Integration

Here, we give a brief account of *multidimensional numerical integration*. For simplicity, we illustrate with the trapezoid rule for the interval $[0, 1]$, using $n + 1$ equally spaced points. The step size is therefore $h = 1/n$. The composite trapezoid rule is then

$$\int_0^1 f(x)\,dx \approx \frac{1}{2h}\left[f(0) + 2\sum_{i=1}^{n-1} f\left(\frac{i}{n}\right) + f(1)\right]$$

We write this in the form

$$\int_0^1 f(x)\,dx \approx \sum_{i=0}^{n} C_i f\left(\frac{i}{n}\right)$$

where

$$C_i = \begin{cases} 1/(2h), & i = 0 \\ 1/h, & 0 < i < n \\ 1/(2h), & i = n \end{cases}$$

The error is $\mathcal{O}(h^2) = \mathcal{O}(n^{-2})$ for functions having a continuous second derivative.

If one is faced with a **two-dimensional integration over the unit square**, then the trapezoid rule can be applied twice:

$$\begin{aligned}\int_0^1\int_0^1 f(x, y)\,dx\,dy &\approx \int_0^1 \sum_{\alpha_1=0}^{n} C_{\alpha_1} f\left(\frac{\alpha_1}{n}, y\right) dy \\ &= \sum_{\alpha_1=0}^{n} C_{\alpha_1} \int_0^1 f\left(\frac{\alpha_1}{n}, y\right) dy \\ &\approx \sum_{\alpha_1=0}^{n} C_{\alpha_1} \sum_{\alpha_2=0}^{n} C_{\alpha_2} f\left(\frac{\alpha_1}{n}, \frac{\alpha_2}{n}\right) \\ &= \sum_{\alpha_1=0}^{n}\sum_{\alpha_2=0}^{n} C_{\alpha_1} C_{\alpha_2} f\left(\frac{\alpha_1}{n}, \frac{\alpha_2}{n}\right)\end{aligned}$$

The error here is again $\mathcal{O}(h^2)$, because each of the two applications of the trapezoid rule entails an error of $\mathcal{O}(h^2)$.

In the same way, we can integrate a function of k variables. Suitable notation is the vector $x = (x_1, x_2, \dots, x_k)^T$ for the independent variable. The region now is taken to be the k-dimensional cube $[0, 1]^k \equiv [0, 1] \times [0, 1] \times \cdots \times [0, 1]$. Then we obtain a **multidimensional numerical integration rule**

$$\int_{[0,1]^k} f(\boldsymbol{x})\,d\boldsymbol{x} \approx \sum_{\alpha_1=0}^{n}\sum_{\alpha_2=0}^{n}\cdots\sum_{\alpha_k=0}^{n} C_{\alpha_1} C_{\alpha_2}\cdots C_{\alpha_k} f\left(\frac{\alpha_1}{n}, \frac{\alpha_2}{n}, \dots, \frac{\alpha_k}{n}\right)$$

The error is still $\mathcal{O}(h^2) = \mathcal{O}(n^{-2})$, provided that f has continuous partial derivatives $\partial^2 f/\partial x_i^2$.

Besides the error involved, one must consider the effort, or work, required to attain a desired level of accuracy. The work in the one-variable case is $\mathcal{O}(n)$. In the two-variable case, it is $\mathcal{O}(n^2)$, and it is $\mathcal{O}(n^k)$ for k variables. The error, now expressed as a function of the number of nodes $N = n^k$, is

$$\mathcal{O}(h^2) = \mathcal{O}(n^{-2}) = \mathcal{O}\Big((n^k)^{-2/k}\Big) = \mathcal{O}(N^{-2/k})$$

Thus, the quality of the numerical approximation of the integral declines very quickly as the number of variables, k, increases. Expressed in other terms, if a constant order of accuracy is to be retained while the number of variables, k, goes up, the number of nodes must go up like n^k. These remarks indicate why the Monte Carlo method for numerical integration becomes more attractive for high-dimensional integration. (This subject is discussed in Chapter 13.)

Summary

(1) To estimate $\int_a^b f(x)\,dx$, divide the interval $[a, b]$ into subintervals according to the partition $P = \{a = x_0 < x_1 < x_2 < \cdots < x_n = b\}$. The **basic trapezoid rule** for the subinterval $[x_i, x_{i+1}]$ is

$$\int_{x_i}^{x_{i+1}} f(x)\,dx \approx A_i = \frac{1}{2}(x_{i+1} - x_i)[f(x_i) + f(x_{i+1})]$$

where the error is $-\frac{1}{12}(x_{i+1} - x_i)^3 f''(\xi_i)$. The **composite trapezoid rule** is

$$\int_a^b f(x)\,dx \approx T(f; P) = \sum_{i=0}^{n-1} A_i = \frac{1}{2}\sum_{i=0}^{n-1}(x_{i+1} - x_i)[f(x_i) + f(x_{i+1})]$$

where the error is $-\frac{1}{12}\sum_{i=1}^{n}(x_{i+1} - x_i)^2 f''(\xi_i)$.

(2) For uniform spacing of nodes in the interval $[a, b]$, we let $x_i = a + ih$, where $h = (b - a)/n$ and $0 \leqq i \leqq n$. The **composite trapezoid rule with uniform spacing** is

$$\int_a^b f(x)\,dx \approx T(f; P) = \frac{h}{2}[f(x_0) + f(x_n)] + h\sum_{i=1}^{n-1} f(x_i)$$

where the error is $-\frac{1}{12}(b - a)^2 f''(\zeta)$.

(3) For uniform spacing of nodes in the interval $[a, b]$ with 2^n subintervals, we let $h = (b - a)/2^n$, and we have

$$\begin{cases} R(0, 0) = \frac{1}{2}(b - a)[f(a) + f(b)] \\ R(n, 0) = h\displaystyle\sum_{i=1}^{2^n - 1} f(a + ih) + \frac{h}{2}[f(a) + f(b)] \end{cases}$$

We can compute the first column of the array $R(n, 0)$ recursively by the **Recursive Trapezoid Formula**:

$$R(n, 0) = \frac{1}{2}R(n - 1, 0) + h\sum_{k=1}^{2^{n-1}} f[a + (2k - 1)h]$$