7.2 Composite Trapezoidal and Simpson's Rule

An intuitive method of finding the area under the curve $y = f(x)$ over $[a, b]$ is by approximating that area with a series of trapezoids that lie above the intervals $\{[x_k, x_{k+1}]\}$.

Theorem 7.2 (Composite Trapezoidal Rule). Suppose that the interval $[a, b]$ is subdivided into M subintervals $[x_k, x_{k+1}]$ of width $h = (b-a)/M$ by using the equally spaced nodes $x_k = a + kh$, for $k = 0, 1, \ldots, M$. The ***composite trapezoidal rule for M subintervals*** can be expressed in any of three equivalent ways:

(1a)
$$T(f, h) = \frac{h}{2}\sum_{k=1}^{M}(f(x_{k-1}) + f(x_k))$$

or

(1b)
$$T(f, h) = \frac{h}{2}(f_0 + 2f_1 + 2f_2 + 2f_3 + \cdots + 2f_{M-2} + 2f_{M-1} + f_M)$$

or

(1c)
$$T(f, h) = \frac{h}{2}(f(a) + f(b)) + h\sum_{k=1}^{M-1} f(x_k).$$

This is an approximation to the integral of $f(x)$ over $[a, b]$, and we write

(2)
$$\int_a^b f(x)\,dx \approx T(f, h).$$

Proof. Apply the trapezoidal rule over each subinterval $[x_{k-1}, x_k]$ (see Figure 7.6). Use the additive property of the integral for subintervals:

(3)
$$\int_a^b f(x)\,dx = \sum_{k=1}^{M}\int_{x_{k-1}}^{x_k} f(x)\,dx \approx \sum_{k=1}^{M}\frac{h}{2}(f(x_{k-1}) + f(x_k)).$$

Since $h/2$ is a constant, the distributive law of addition can be applied to obtain (1a). Formula (1b) is the expanded version of (1a). Formula (1c) shows how to group all the intermediate terms in (1b) that are multiplied by 2. •

Approximating $f(x) = 2 + \sin(2\sqrt{x})$ with piecewise linear polynomials results in places where the approximation is close and places where it is not. To achieve accuracy, the composite trapezoidal rule must be applied with many subintervals. In the next example we have chosen to integrate this function numerically over the interval [1, 6]. Investigation of the integral over [0, 1] is left as an exercise.

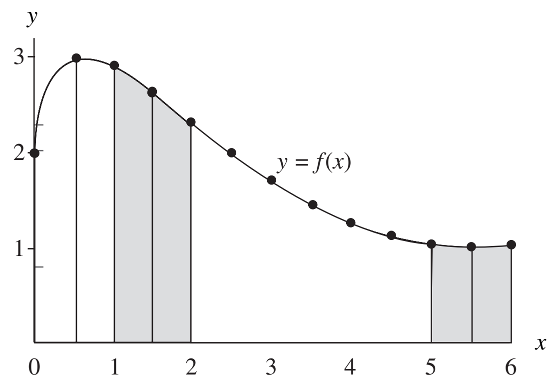

Figure 7.6 Approximating the area under the curve $y = 2 + \sin(2\sqrt{x})$ with the composite trapezoidal rule.

Example 7.5. Consider $f(x) = 2 + \sin(2\sqrt{x})$. Use the composite trapezoidal rule with 11 sample points to compute an approximation to the integral of $f(x)$ taken over $[1, 6]$.

To generate 11 sample points, we use $M = 10$ and $h = (6 - 1)/10 = 1/2$. Using formula (1c), the computation is

$$
\begin{aligned}
T(f, \tfrac{1}{2}) &= \frac{1/2}{2}(f(1) + f(6)) \\
&+ \frac{1}{2}(f(\tfrac{3}{2}) + f(2) + f(\tfrac{5}{2}) + f(3) + f(\tfrac{7}{2}) + f(4) + f(\tfrac{9}{2}) + f(5) + f(\tfrac{11}{2})) \\
&= \frac{1}{4}(2.90929743 + 1.01735756) \\
&+ \frac{1}{2}(2.63815764 + 2.30807174 + 1.97931647 + 1.68305284 + 1.43530410 \\
&+ 1.24319750 + 1.10831775 + 1.02872220 + 1.00024140) \\
&= \frac{1}{4}(3.92665499) + \frac{1}{2}(14.42438165) \\
&= 0.98166375 + 7.21219083 = 8.19385457.
\end{aligned}
$$

■

Error Analysis

The significance of the next two results is to understand that the error terms $E_T(f, h)$ and $E_S(f, h)$ for the composite trapezoidal rule and composite Simpson rule are of the order $\boldsymbol{O}(h^2)$ and $\boldsymbol{O}(h^4)$, respectively. This shows that the error for Simpson's rule converges to zero faster than the error for the trapezoidal rule as the step size h decreases to zero. In cases where the derivatives of $f(x)$ are known, the formulas

$$E_T(f, h) = \frac{-(b-a)f^{(2)}(c)h^2}{12} \quad \text{and} \quad E_S(f, h) = \frac{-(b-a)f^{(4)}(c)h^4}{180}$$

can be used to estimate the number of subintervals required to achieve a specified accuracy.

Corollary 7.2 (Trapezoidal Rule: Error Analysis). Suppose that $[a, b]$ is subdivided into M subintervals $[x_k, x_{k+1}]$ of width $h = (b-a)/M$. The composite trapezoidal rule

$$T(f, h) = \frac{h}{2}(f(a) + f(b)) + h \sum_{k=1}^{M-1} f(x_k) \tag{7}$$

is an approximation to the integral

$$\int_a^b f(x)\,dx = T(f, h) + E_T(f, h). \tag{8}$$

Furthermore, if $f \in C^2[a, b]$, there exists a value c with $a < c < b$ so that the error term $E_T(f, h)$ has the form

$$E_T(f, h) = \frac{-(b-a)f^{(2)}(c)h^2}{12} = \boldsymbol{O}(h^2). \tag{9}$$

Proof. We first determine the error term when the rule is applied over $[x_0, x_1]$. Integrating the Lagrange polynomial $P_1(x)$ and its remainder yields

$$\int_{x_0}^{x_1} f(x)\,dx = \int_{x_0}^{x_1} P_1(x)\,dx + \int_{x_0}^{x_1} \frac{(x-x_0)(x-x_1)f^{(2)}(c(x))}{2!}\,dx. \tag{10}$$

The term $(x - x_0)(x - x_1)$ does not change sign on $[x_0, x_1]$, and $f^{(2)}(c(x))$ is continuous. Hence the second mean value theorem for integrals implies that there exists a value c_1 so that

$$\int_{x_0}^{x_1} f(x)\,dx = \frac{h}{2}(f_0 + f_1) + f^{(2)}(c_1)\int_{x_0}^{x_1} \frac{(x-x_0)(x-x_1)}{2!}\,dx. \tag{11}$$

Use the change of variable $x = x_0 + ht$ in the integral on the right side of (11):

$$\begin{aligned}\int_{x_0}^{x_1} f(x)\,dx &= \frac{h}{2}(f_0 + f_1) + \frac{f^{(2)}(c_1)}{2}\int_0^1 h(t-0)h(t-1)h\,dt\\ &= \frac{h}{2}(f_0 + f_1) + \frac{f^{(2)}(c_1)h^3}{2}\int_0^1 (t^2 - t)\,dt\\ &= \frac{h}{2}(f_0 + f_1) - \frac{f^{(2)}(c_1)h^3}{12}.\end{aligned} \tag{12}$$

Now we are ready to add up the error terms for all of the intervals $[x_k, x_{k+1}]$:

$$\begin{aligned}\int_a^b f(x)\,dx &= \sum_{k=1}^{M}\int_{x_{k-1}}^{x_k} f(x)\,dx\\ &= \sum_{k=1}^{M}\frac{h}{2}(f(x_{k-1}) + f(x_k)) - \frac{h^3}{12}\sum_{k=1}^{M} f^{(2)}(c_k).\end{aligned} \tag{13}$$

The first sum is the composite trapezoidal rule $T(f, h)$. In the second term, one factor of h is replaced with its equivalent $h = (b - a)/M$, and the result is

$$\int_a^b f(x)\,dx = T(f, h) - \frac{(b-a)h^2}{12}\left(\frac{1}{M}\sum_{k=1}^{M} f^{(2)}(c_k)\right).$$

The term in parentheses can be recognized as an average of values for the second derivative and hence is replaced by $f^{(2)}(c)$. Therefore, we have established that

$$\int_a^b f(x)\,dx = T(f, h) - \frac{(b-a)f^{(2)}(c)h^2}{12},$$

and the proof of Corollary 7.2 is complete. •

Example 7.7. Consider $f(x) = 2 + \sin(2\sqrt{x})$. Investigate the error when the composite trapezoidal rule is used over $[1, 6]$ and the number of subintervals is 10, 20, 40, 80, and 160.

Table 7.2 shows the approximations $T(f, h)$. The antiderivative of $f(x)$ is

$$F(x) = 2x - \sqrt{x}\cos(2\sqrt{x}) + \frac{\sin(2\sqrt{x})}{2},$$

and the true value of the definite integral is

$$\int_1^6 f(x)\,dx = F(x)\Big|_{x=1}^{x=6} = 8.1834792077.$$

This value was used to compute the values $E_T(f, h) = 8.1834792077 - T(f, h)$ in Table 7.2. It is important to observe that when h is reduced by a factor of $\frac{1}{2}$ the successive errors $E_T(f, h)$ are diminished by approximately $\frac{1}{4}$. This confirms that the order is $\boldsymbol{O}(h^2)$. ■

Table 7.2 Composite Trapezoidal Rule for $f(x) = 2 + \sin(2\sqrt{x})$ over [1, 6]

M	h	$T(f, h)$	$E_T(f, h) = \boldsymbol{O}(h^2)$
10	0.5	8.19385457	−0.01037540
20	0.25	8.18604926	−0.00257006
40	0.125	8.18412019	−0.00064098
80	0.0625	8.18363936	−0.00016015
160	0.03125	8.18351924	−0.00004003

Example 7.9. Find the number M and the step size h so that the error $E_T(f, h)$ for the composite trapezoidal rule is less than 5×10^{-9} for the approximation $\int_2^7 dx/x \approx T(f, h)$.

The integrand is $f(x) = 1/x$ and its first two derivatives are $f'(x) = -1/x^2$ and $f^{(2)}(x) = 2/x^3$. The maximum value of $|f^{(2)}(x)|$ taken over [2, 7] occurs at the endpoint $x = 2$, and thus we have the bound $|f^{(2)}(c)| \le |f^{(2)}(2)| = \frac{1}{4}$, for $2 \le c \le 7$. This is used with formula (9) to obtain

$$(17) \qquad |E_T(f, h)| = \frac{|-(b-a)f^{(2)}(c)h^2|}{12} \le \frac{(7-2)\frac{1}{4}h^2}{12} = \frac{5h^2}{48}.$$

The step size h and number M satisfy the relation $h = 5/M$, and this is used in (17) to get the relation

$$(18) \qquad |E_T(f, h)| \le \frac{125}{48M^2} \le 5 \times 10^{-9}.$$

Now rewrite (18) so that it is easier to solve for M:

$$(19) \qquad \frac{25}{48} \times 10^9 \le M^2.$$

Solving (19), we find that $22821.77 \le M$. Since M must be an integer, we choose $M = 22{,}822$, and the corresponding step size is $h = 5/22{,}822 = 0.000219086846$. When the composite trapezoidal rule is implemented with this many function evaluations, there is a possibility that the rounded-off function evaluations will produce a significant amount of error. When the computation was performed, the result was

$$T\left(f, \frac{5}{22{,}822}\right) = 1.252762969,$$

which compares favorably with the true value $\int_2^7 dx/x = \ln(x)|_{x=2}^{x=7} = 1.252762968$. The error is smaller than predicted because the bound $\frac{1}{4}$ for $|f^{(2)}(c)|$ was used. Experimentation shows that it takes about 10,001 function evaluations to achieve the desired accuracy of 5×10^{-9}, and when the calculation is performed with $M = 10{,}000$, the result is

$$T\left(f, \frac{5}{10{,}000}\right) = 1.252762973. \qquad \blacksquare$$

The composite trapezoidal rule usually requires a large number of function evaluations to achieve an accurate answer. This is contrasted in the next example with Simpson's rule, which will require significantly fewer evaluations.

Numerical Methods Using Matlab, 4th Edition, 2004

John H. Mathews and Kurtis K. Fink

ISBN: 0-13-065248-2

Prentice-Hall Inc.
Upper Saddle River, New Jersey, USA
http://vig.prenhall.com/