8. Approximation of Integrals

Basic idea: replace the function by its interpolating polynomial and use the integral of the interpolating polynomial as an approximation to the integral of the function.

8.1. **Basic Numerical Integration Rules.**

Let P_n be the polynomial of degree $\le n$ interpolating f at $x_0, \dots, x_n$. Then $f(x) = P_n(x) + f[x_0, \dots, x_n, x]\psi_n(x)$, where $\psi_n(x) = \prod_{j=0}^n (x - x_j)$. Hence,

$$\int_a^b f(x)\,dx = \int_a^b P_n(x)\,dx + \int_a^b f[x_0, \dots, x_n, x]\psi_n(x)\,dx.$$

To simplify the error formula, we can use the Mean Value Theorem for integrals, i.e.,

Theorem 4. *Let $g(x)$ be integrable and of one sign on $[a, b]$. If $F(x)$ is continuous on $[a, b]$, then*

$$\int_a^b F(x)g(x)\,dx = F(\xi)\int_a^b g(x)\,dx, \qquad \textit{for some } \xi \in [a, b].$$

Applying this theorem, we get that if $\psi_n(x)$ is of one sign on (a, b), then

$$\int_a^b f(x)\,dx = \int_a^b P_n(x)\,dx + f[x_0, \dots, x_n, \xi]\int_a^b \psi_n(x)\,dx.$$

Furthermore, if $f \in C^{n+1}$, then

$$E(f) \equiv \int_a^b [f(x) - P_n(x)]\,dx = \frac{1}{(n+1)!} f^{(n+1)}(\eta)\int_a^b \psi_n(x)\,dx.$$

There is then a further simplification, since $\psi_n(x)$ can be integrated exactly.

Another case when the error formula can be simplified is when $\int_a^b \psi_n(x)\,dx = 0$. Noting that

$$\begin{aligned} f[x_0, \dots, x_n, x_{n+1}, x] = f[x_{n+1}, x_0, \dots, x_n, x] &= \frac{f[x_0, \dots, x_n, x] - f[x_{n+1}, x_0, \dots, x_n]}{x - x_{n+1}} \\ &= \frac{f[x_0, \dots, x_n, x] - f[x_0, \dots, x_n, x_{n+1}]}{x - x_{n+1}}, \end{aligned}$$

we get the identity:

$$f[x_0, \dots, x_n, x] = f[x_0, \dots, x_n, x_{n+1}] + f[x_0, \dots, x_n, x_{n+1}, x](x - x_{n+1}).$$

Hence, if $\int_a^b \psi_n(x)\,dx = 0$,

$$\begin{aligned} E(f) = \int_a^b f[x_0, \dots, x_n, x]\psi_n(x)\,dx &= f[x_0, \dots, x_n, x_{n+1}]\int_a^b \psi_n(x)\,dx \\ + \int_a^b f[x_0, \dots, x_n, x_{n+1}, x](x - x_{n+1})\psi_n(x)\,dx &= \int_a^b f[x_0, \dots, x_n, x_{n+1}, x](x - x_{n+1})\psi_n(x)\,dx. \end{aligned}$$

If we can choose $x_{n+1} \in [a, b]$ so that $(x - x_{n+1})\psi_n(x)$ is of one sign in (a, b), then we will get

$$\begin{aligned} E(f) = f[x_0, \ldots, x_n, x_{n+1}, \xi] \int_a^b (x - x_{n+1})\psi_n(x)\,dx \\ = \frac{1}{(n+2)!} f^{(n+2)}(\eta) \int_a^b (x - x_{n+1})\psi_{n+1}(x)\,dx, \end{aligned}$$

for some $\eta \in [a, b]$.

Examples:

$\mathbf{n = 0}$. $f(x) = f(x_0) + f[x_0, x](x - x_0)$. Then

$$\int_a^b f(x)\,dx = (b - a)f(x_0) + \int_a^b f[x_0, x](x - x_0)\,dx.$$

If $x_0 = a$, then $\psi_0(x) = x - a$ is of one sign, so

$$\int_a^b f(x)\,dx = (b - a)f(a) + f'(\eta) \int_a^b (x - a)\,dx = (b - a)f(a) + f'(\eta)(b - a)^2/2.$$

If $x_0 = b$, then $\psi_0(x) = x - b$ is of one sign, so

$$\int_a^b f(x)\,dx = (b - a)f(b) + f'(\eta) \int_a^b (x - b)\,dx = (b - a)f(b) - f'(\eta)(b - a)^2/2.$$

These formulas are known as rectangle rules.

If $x_0 = (a + b)/2$, then $\psi_0(x)$ is not of one sign on (a, b), but $\int_a^b \psi_0(x)\,dx = 0$. If we choose $x_1 = x_0$, then $\psi_1(x) = (x - x_0)^2$ is of one sign, so we obtain

$$\begin{aligned} \int_a^b f(x)\,dx = (b - a)f([a + b]/2) + \frac{1}{2!} f''(\eta) \int_a^b [x - (a + b)/2]^2\,dx \\ = (b - a)f([a + b]/2) + \frac{1}{24} f''(\eta)(b - a)^3. \end{aligned}$$

This formula is called the midpoint rule.

$\mathbf{n = 1}$. $f(x) = f(x_0) + f[x_0, x_1](x - x_0) + f[x_0, x_1, x](x - x_0)(x - x_1)$. If we choose $x_0 = a$ and $x_1 = b$, then $\psi_1(x) = (x - a)(x - b)$ is always of one sign on (a, b), so we get

$$\begin{aligned} \int_a^b f(x)\,dx &= \int_a^b \{f(a) + f[a, b](x - a)\}\,dx + \frac{1}{2!} f''(\eta) \int_a^b (x - a)(x - b)\,dx \\ &= f(a)(b - a) + f[a, b](b - a)^2/2 - \frac{1}{12} f''(\eta)(b - a)^3 \\ &= \frac{b - a}{2}[f(a) + f(b)] - \frac{1}{12} f''(\eta)(b - a)^3. \end{aligned}$$

Formula called Trapezoidal Rule.

n = 2. $f(x) = f(x_0) + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1) + f[x_0, x_1, x_2, x](x - x_0)(x - x_1)(x - x_2)$. Now for x_0, x_1, x_2 distinct points in $[a, b]$, $\psi_2(x) = (x - x_0)(x - x_1)(x - x_2)$ is not of one sign in $[a, b]$. But if $x_0 = a$, $x_1 = (a + b)/2$, $x_2 = b$, then $\int_a^b \psi_2(x)\,dx = 0$. If we choose $x_3 = x_1$, then $\psi_3(x) = (x - a)(x - [a + b]/2)^2(x - b)$, which is of one sign in $[a, b]$. Hence, we get the error formula

$$\int_a^b f(x)\,dx - \int_a^b P_2(x)\,dx = -\frac{1}{90}\left(\frac{b-a}{2}\right)^5 f^{(4)}(\eta).$$

The resulting quadrature formula is:

$$\int_a^b f(x)\,dx \approx \frac{b-a}{6}[f(a) + 4f([a+b]/2) + f(b)], \qquad \text{Simpson's rule.}$$

In general, formulas obtained by integrating interpolation points using equally spaced interpolation points are called Newton-Cotes quadrature formulas. There are two types: (1) closed formulas in which the end points of the interval are used in the integration formula and (2) open formulas in which the end points of the interval are not used and the other points are symmetrically placed.

8.2. **Composite Numerical Integration Rules.** In practice, we use composite formulas based on integration of piecewise polynomials, i.e., we subdivide the interval $[a, b]$ into subintervals $[x_{i-1}, x_i]$, where $a = x_0 < x_1 < \ldots < x_N = b$ and use the fact that $\int_a^b f(x)\,dx = \sum_{i=1}^N \int_{x_{i-1}}^{x_i} f(x)\,dx$. We then approximate each of integrals $\int_{x_{i-1}}^{x_i} f(x)\,dx$ by one of the formulas just developed and add the results.

Consider the case when the interpolation points are equally spaced, i.e., $x_i = a + ih$, $i = 0, \ldots N$, so that $h = (b - a)/N$. Set $f_s = f(a + sh)$ so that $f_i = f(x_i)$. We then obtain the following composite quadrature formulas.

Example: Composite Midpoint rule

$$\int_{x_{i-1}}^{x_i} f(x)\,dx = hf_{i-1/2} + h^3 f^{(2)}(\xi_i)/24, \quad \xi_i \in [x_{i-1}, x_i]$$

so that

$$\int_a^b f(x)\,dx = h\sum_{i=1}^N f_{i-1/2} + \frac{h^3}{24}\sum_{i=1}^N f^{(2)}(\xi_i).$$

The error term can be simplified by using the mean value theorem for sums:

Theorem 5. *Let $F(x)$ be a continuous function on $[a, b]$, let $x_1, \cdots, x_n$ be points in $[a, b]$, and let $g_1, \cdots, g_n$ be real numbers all of one sign. Then $\sum_{i=1}^n F(x_i)g_i = F(\xi)\sum_{i=1}^n g_i$ for some $\xi \in [a, b]$.*

Choosing $F(x) = f^{(2)}(x)$, $x_i = \xi_i$, $g_i = h^3/24$, and using the fact that $Nh = b - a$, we get that if $f^{(2)}$ is continuous on $[a, b]$, then

$$\int_a^b f(x)\,dx - h\sum_{i=1}^{N} f_{i-1/2} = \frac{h^3}{24} N f^{(2)}(\xi) = \frac{b-a}{24} h^2 f^{(2)}(\xi).$$

Note that if $|f^{(2)}(x)| \le M$ for all $x \in [a, b]$, then by choosing h sufficiently small, we can achieve any desired accuracy (neglecting roundoff error in the computation).

Composite Trapezoidal rule:

$$\int_{x_{i-1}}^{x_i} f(x)\,dx = \frac{h}{2}[f_{i-1} + f_i] - h^3 f^{(2)}(\xi_i)/12, \quad \xi_i \in [x_{i-1}, x_i]$$

so that

$$\begin{aligned}\int_a^b f(x)\,dx &= \frac{h}{2}\sum_{i=1}^{N}[f_{i-1} + f_i] - \frac{h^3}{12}\sum_{i=1}^{N} f^{(2)}(\xi_i)\\ &= h[\frac{1}{2}f_0 + f_1 + f_2 + \cdots + f_{N-1} + \frac{1}{2}f_N] - \frac{b-a}{12}h^2 f^{(2)}(\xi).\end{aligned}$$

Composite Simpson's rule: (on N subintervals)

$$\int_{x_{i-1}}^{x_i} f(x)\,dx = \frac{h}{6}[f_{i-1} + 4f_{i-1/2} + f_i] - \frac{1}{90}\left(\frac{h}{2}\right)^5 f^{(4)}(\xi_i), \quad \xi_i \in [x_{i-1}, x_i]$$

so that

$$\begin{aligned}\int_a^b f(x)\,dx &= \frac{h}{6}\sum_{i=1}^{N}[f_{i-1} + 4f_{i-1/2} + f_i] - \frac{1}{90}\left(\frac{h}{2}\right)^5 \sum_{i=1}^{N} f^{(4)}(\xi_i)\\ &= \frac{h}{6}[f_0 + f_N + 2\sum_{i=1}^{N-1} f_i + 4\sum_{i=1}^{N} f_{i-1/2}] - \frac{b-a}{180}\left(\frac{h}{2}\right)^4 f^{(4)}(\xi).\end{aligned}$$

If an upper bound on the proper derivative is known on the interval $[a, b]$, then the error formula could be used to determine a value of h (or equivalently the number of subintervals) that would guarantee any desired accuracy. Since often a bound is not known, or even if known is a worst case estimate, one does not usually precalculate the number of subintervals needed to guarantee a given accuracy. In the next sections, we shall consider more computationally efficient approaches to finding accurate approximations to the integral.