Lecture 19: Numerical Integration II

Instructor: Professor Amos Ron Scribes: Mark Cowlishaw, Nathanael Fillmore

1 Error Analysis of Simple Rules for Numerical Integration

Last time we discussed approximating the definite integral

$$I(f) \quad = \quad \int_a^b f(t)dt$$

The general approach introduced last time was to interpolate function f using some polynomial $p(t)$, choosing interpolation points according to some rule r and compute the integral of the polynomial, $\int_a^b p(t)dt$ as the approximation. Let $X = \{t_0, t_1, \cdots, t_n\}$ be the interpolating node set. We can write the integral of $p(t)$ using Lagrange polynomials:

$$\begin{aligned} I_r(f) &= \int_a^b p(t)dt \\ &= \int_a^b \left(f(t_0)\ell_0(t) + f(t_1)\ell_1(t) + \cdots + f(t_n)\ell_n(t)\right) dt \\ &= \int_a^b \sum_{i=0}^n f(t_i)\ell_i(t)dt \\ &= \sum_{i=0}^n f(t_i) \underbrace{\int_a^b \ell_i(t)dt}_{w_i} \end{aligned}$$

Since the Lagrange polynomials $\ell_i(t)$ depend only on the interpolation points and not the corresponding function values, we can rewrite this approximation as a simple weighted sum of function values:

$$I_r(f) \quad = \quad \sum_{i=0}^n f(t_i)w_i$$

Last time we presented four rules that used this scheme to approximate a definite integral:

Rectangle Rule

The rectangle rule uses node set $X = \{a\}$, the left endpoint of the interval $[a, b]$ to interpolate $f|_{[a,b]}$ using a constant polynomial ($p(t) = f(a)$). The corresponding estimate of the definite integral is given by:

$$I_R \quad = \quad f(a)(b - a)$$

Midpoint Rule

The midpoint rule uses node set $X = \left\{\frac{a+b}{2}\right\}$, the midpoint of the interval $[a, b]$ to interpolate $f|_{[a,b]}$ using a constant polynomial ($p(t) = f(\frac{a+b}{2})$). The corresponding estimate of the definite integral is given by:

$$I_M = f\left(\frac{a+b}{2}\right)(b-a)$$

Trapezoid Rule

The trapezoid rule uses node set $X = \{a, b\}$, the left and right endpoints of the interval $[a, b]$ to interpolate $f|_{[a,b]}$ using a polynomial of degree at most 1 ($p(t) = f(a)\frac{t-b}{a-b} + f(b)\frac{t-a}{b-a}$). The corresponding estimate of the definite integral is given by:

$$I_T = (f(a) + f(b))\frac{b-a}{2}$$

Simpson's Rule

Simpson's rule uses node set $X = \left\{a, \frac{a+b}{2}, b\right\}$, the left endpoint, midpoint, and right endpoint of the interval $[a, b]$ to interpolate $f|_{[a,b]}$ using a polynomial of degree at most 2 ($p(t) = f(a)\frac{(t-b)(t-m)}{(a-b)(a-m)} + f(m)\frac{(t-a)(t-b)}{(m-a)(m-b)} + f(b)\frac{(t-a)(t-m)}{(b-a)(b-m)}$, where m is the midpoint of $[a, b]$). The corresponding estimate of the definite integral is given by:

$$I_S = \left(f(a) + 4f\left(\frac{a+b}{2}\right) + f(b)\right)\frac{b-a}{6}$$

In last lecture's example, we estimated $\ln(1.2)$ using the four rules and obtained the following results:

$$\begin{aligned} I_R &= 0.2 \\ I_M &= 0.181818\cdots \\ I_T &= 0.183333\cdots \\ I_S &= 0.182323\cdots \end{aligned}$$

1.1 Error Analysis

Recall that last time we showed that the error of approximating a definite integral using polynomial interpolation over $T = \{t_0, t_1, \ldots, t_n\}$ is given by:

$$\begin{aligned} E_r(f) &= \int_a^b [f(t) - p(t)]\, dt \\ &= \int_a^b \left[\frac{f^{(n+1))}(c)}{(n+1)!}\underbrace{\prod_{i=0}^{n}(t-t_i)}_{\omega(t)}\right] dt \end{aligned}$$

We split the error analysis into two cases:

Case 1: $\omega(t)|_{[a,b]}$ is always nonnegative, or always non-positive

In this case, we can calculate the error as:

$$E_r(f) = \frac{f^{(n+1)}(c)}{(n+1)!}\int_a^b \omega(t)dt$$

The Rectangle and Trapezoid rule fit this case, and last time we showed that the error for each can be written as:

$$\begin{aligned} E_R(f) &= \frac{f'(c)}{2}(b-a)^2 \\ E_T(f) &= \frac{f''(c)}{12}(b-a)^3 \end{aligned}$$

Case 2: $\int_a^b \omega(t)dt = 0$

It is easy to see that the midpoint rule falls into this case, since:

$$\begin{aligned} \int_a^b \omega_M(t)dt &= \int_a^b \left(t - \frac{a+b}{2}\right) dt \\ &= \left.\frac{[t-(a+b)/2]^2}{2}\right|_a^b \\ &= 0 \end{aligned}$$

and Simpson's rule behaves similarly. An interesting property of rules that fall into case 2 is that adding another interpolation point does not change the integral of the polynomial interpolant. This is easy to see, since $\omega(t)$ is the next Newton polynomial and since its integral is 0, the weight of the corresponding function value w_{n+1} will be 0.

1.2 Error Analysis of Midpoint Rule

Since the midpoint rule fits into case 2 of our error analysis, that is:

$$\begin{aligned} \int_a^b \omega(t)dt &= \left.\frac{[t-(a+b)/2]^2}{2}\right|_a^b \\ &= 0 \end{aligned}$$

as shown in Figure 1, we can add an interpolation point without affecting the area of the interpolated polynomial, leaving the error unchanged. We can therefore do our error analysis of the midpoint rule with any single point added - since adding *any* point in $[a, b]$ does not affect the area, we simply double the midpoint, so that $X = \{(a+b)/2, (a+b)/2\}$. We can now examine the value of the next Newton polynomial, $\omega(t)$ for the modified rule:

$$\omega(t) = \left(t - \frac{a+b}{2}\right)\left(t - \frac{a+b}{2}\right)$$

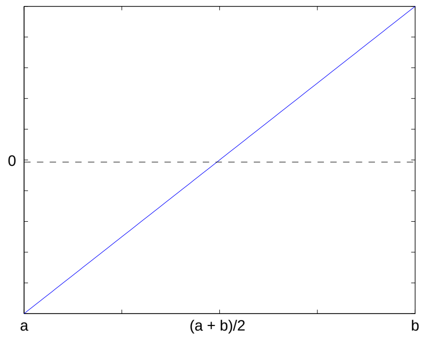

Figure 1: $\omega(t)$ in the Midpoint rule over $[a, b]$

Clearly, $\omega(t)|_{[a,b]} \geq 0$, so that this new rule can be analyzed using case 1, this yields:

$$\begin{aligned}
E_M(f) &= \frac{f''(c)}{2} \int_a^b \left(t - \frac{a+b}{2}\right)^2 dt \\
&= \frac{f''(c)}{2} \left. \frac{\left(t - \frac{a+b}{2}\right)^3}{3} \right|_a^b \\
&= \frac{f''(c)}{2} \frac{\left(\frac{b-a}{2}\right)^3 - \left(\frac{a-b}{2}\right)^3}{3} \\
&= \frac{f''(c)}{2} \frac{2(b-a)^3}{24} \\
&= \frac{f''(c)}{24} (b-a)^3
\end{aligned}$$

Note that this error is a constant factor of two smaller than the error for the trapezoid rule.

1.3 Error Analysis of Simpson's Rule

Since Simpson's rule also fits into case 2 of our error analysis, that is:

$$\int_a^b \omega(t) dt = 0$$

as shown in Figure 2, we can add an interpolation point without affecting the area of the interpolated polynomial, leaving the error unchanged. We can therefore do our error analysis of Simpson's rule with any single point added - since adding *any* point in $[a, b]$ does not affect the area, we simply double the midpoint, so that our node set $X = \{a, (a+b)/2, (a+b)/2, b\}$. We can now examine the value of the next Newton polynomial, $\omega(t)$ for the modified rule:

$$\omega(t) = (t-a)\left(t - \frac{a+b}{2}\right)^2 (t-b)$$

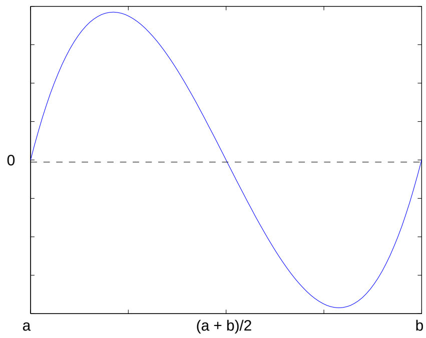

Figure 2: $\omega(t)$ in Simpson's rule over $[a, b]$

Clearly, $\omega(t)|_{[a,b]} \leq 0$, so that this new rule can be analyzed using case 1, this yields:

$$\begin{aligned} E_M(f) &= \frac{f^{(4)}(c)}{24} \int_a^b (t-a)\left(t - \frac{a+b}{2}\right)^2 (t-b)dt \\ &= -\frac{f^{(4)}(c)}{2880}(b-a)^5 \end{aligned}$$

2 Composite Rules

Notice that the error formula for each of the simple rules depends on a high power of the size of the interval $b - a$, so that a small interval makes for a smaller error. This motivates the following general idea for creating composite rules for numerical integration.

Step 1

Partition the interval $[a, b]$ into N subintervals, equidistant by default, with width

$$h = \frac{b-a}{N}$$

Step 2

Apply a simple approximation rule r to each subinterval $[x_i, x_{i+1}]$ and use the area I_r as the approximation of the integral for that subinterval:

$$\int_{x_i}^{x_{i+1}} f(t)dt \approx I_{r\ [x_i, x_{i+1}]}(f)$$

Note that in this application, the appearance of each of the piecewise polynomials is unimportant, we are only interested in their approximation of the definite integral.

Step 3

Add up the approximation of the area over each subinterval to obtain the approximation over the entire interval $[a, b]$:

$$I_{[a,b]}(f) \approx \sum_{i=0}^{n-1} I_{r\ [x_i,x_{i+1}]}(f)$$

EXAMPLE 2.1. *To illustrate, consider applying the composite rectangle rule to an interval $[a, b]$, as shown in Figure 4. In each subinterval, the left endpoint gets weight h. Thus every point except the last one in our partition has weight 1; the last point has weight 0. This yields the following estimate of the definite integral:*

$$I_{CR} = h \cdot f(a) + h \cdot f(a + h) + h \cdot f(a + 2h) + \cdots + h \cdot f(b - h)$$

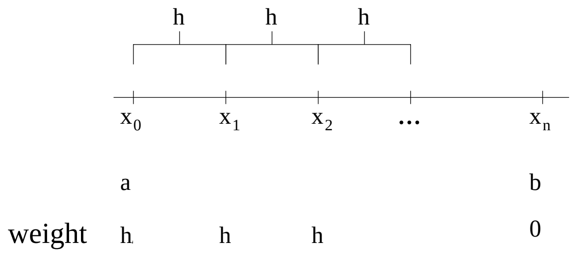

Figure 3: Function Value Weights in the Composite Rectangle Rule

EXAMPLE 2.2. *AS another illustration, consider applying the composite trapezoid rule to an interval $[a, b]$, as shown in Figure 4. In each subinterval, the endpoints get weight $h/2$. Since each of the interior points is included in two subintervals, this yields the following estimate of the definite integral:*

$$I_{CT} = \frac{h}{2}\left[f(a) + 2 \cdot f(a + h) + 2 \cdot f(a + 2h) + \cdots + 2 \cdot f(b - h) + f(b)\right]$$

EXAMPLE 2.3. *Now, consider applying composite Simpson's rule to an interval $[a, b]$ as shown in Figure 5. For each subinterval $[x_i, x_{i+1}]$, the endpoints get weight $1/6$ and the midpoint gets weight $4/6$. Since each interior endpoint (all nodes except a and b) is counted twice, this yields the following estimate of the definite integral:*

$$\begin{aligned} I_{CS} = \frac{h}{6}\Big[& f(a) + 4 \cdot f(a + \frac{h}{2}) + 2 \cdot f(a + h) + 4 \cdot f(a + \frac{3h}{2}) + 2 \cdot f(a + 2h) + \ldots \\ & + 2 \cdot f(b - h) + 4 \cdot f(b - \frac{h}{2}) + f(b)\Big] \end{aligned}$$

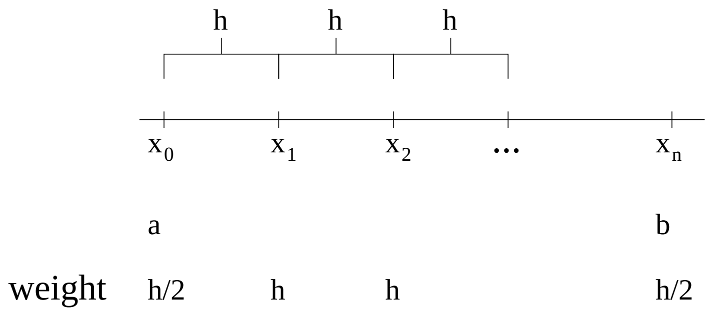

Figure 4: Function Value Weights in the Composite Trapezoid Rule

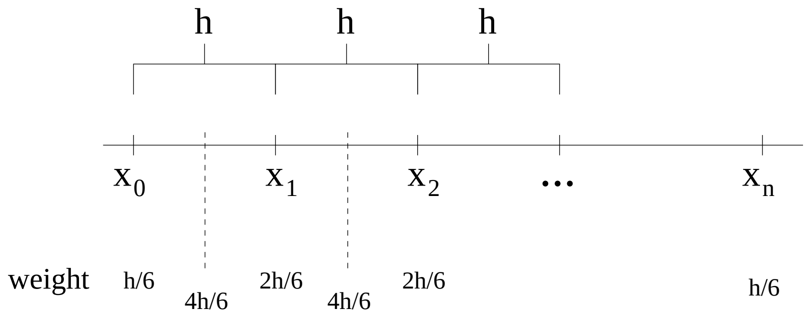

Figure 5: Function Value Weights in Composite Simpson's Rule

2.1 Error Analysis for Composite Simpson's Rule

In a composite rule, we are making use of the fact that a definite integral over an interval $[a, b]$ is simply the sum of the definite integrals of the subintervals.

$$\int_a^b f(t)dt = \sum_{i=0}^{n-1} \int_{x_i}^{x_{i+1}} f(t)dt$$

To get an expression for the error of a composite rule, we take the sum of the errors over each subinterval, noting that over and underestimates may cancel out:

$$\begin{aligned} E_{CS}(f) &= \sum_{i=0}^{n-1} E_{s\ [x_i, x_{i+1}]} \\ &= -\frac{1 \cdot h^5}{2880} \cdot N \cdot \frac{\sum_{i=0}^{n-1} f^{(4)}(c_i)}{N} \\ &= -\frac{1}{2880} \cdot \underbrace{f^{(4)}(c)}_{(*)} \cdot h^4 \cdot (N \cdot h) \\ &= -\frac{f^{(4)}(c)}{2880} h^4 (b-a) \end{aligned}$$

Note that $(*)$ is an average of the values of $f(c_i)$. This is a general formula for the error of a composite rule. In general, a simple rule r with error of the form:

$$E_r = \frac{f^{(n)}(c)}{k}(b-a)^{n+1}$$

will produce a composite rule Cr with error of the form:

$$E_{Cr} = \frac{f^{(n)}(c)}{k}h^n(b-a)$$

We will discuss this further in the next lecture.