Numerical Methods

DS288 and UMC201

Ratikanta Behera

Department of Computational and Data Sciences,
Indian Institute of Science Bangalore

August-December 2025

Composite Trapezoidal rule

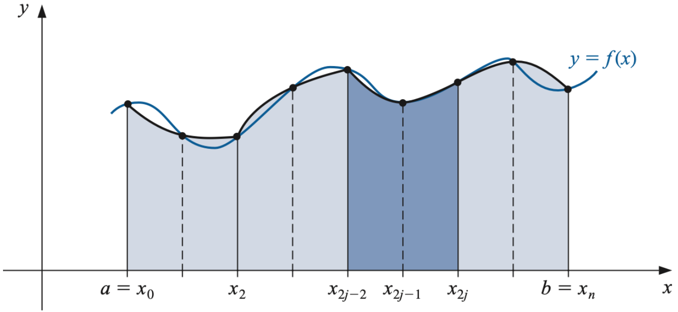
- Let $f \in C^2[a,b]$, $h = (b-a)/n$, and $x_j = a + jh$, for each $j = 0, 1, \cdots, n$.
- There exists $\mu \in (a,b)$ for which the Composite Trapezoidal rule for n sub-intervals can be written with its error term as

$$\int_a^b f(x)dx = \frac{h}{2}\left[f(a) + 2\sum_{j=1}^{n-1} f(x_j) + f(b)\right] - \frac{b-a}{12}h^2 f''(\mu)$$

Composite Simpson's rule

- Let $f \in C^4[a, b]$, n be even, $h = (b - a)/n$, and $x_j = a + jh$, for each $j = 0, 1, \cdots, n$.
- There exists $\zeta \in (a, b)$ for which the Composite Simpson's rule for n sub-intervals can be written with its error term as

$$\int_a^b f(x)dx = \frac{h}{3}\left[f(a) + 2\sum_{j=1}^{(n/2)-1} f(x_{2j}) + 4\sum_{j=1}^{n/2} + f(x_{2j-1}) + f(b)\right] - \frac{b-a}{180}h^4 f^{(4)}(\zeta)$$



Example

Example

Determine values of h that will ensure approximation error of less than 0.00002 while approximating $\int_0^\pi \sin(x)dx$ on $[0, \pi]$ using (a). Composite Trapezoidal rule and (b): Composite Simpson's rule.

(a): Composite Trapezoidal rule

- Error: $\left|\frac{\pi h^2}{12} f''(\mu)\right| = \frac{\pi h^2}{12}|\sin(\mu)| \leq \frac{\pi h^2}{12} < 0.00002$
- since $h = \pi/n$. Thus $n = \pi/h$, we need $\frac{\pi^3}{12n^2} < 0.00002$
- $n \geq \left(\frac{\pi^3}{12(0.00002)}\right)^{1/2} = 359.44 \approx 360$

(b): Composite Simpson's rule

- Error: $|\frac{\pi h^4}{180} f^{(4)}(\mu)| = \frac{\pi h^4}{180}|\sin(\mu)| \leq \frac{\pi h^4}{180} < 0.00002$
- $h \leq 0.18398$, $n \geq 18$

Best Choice for h

- Approximation error is of the form ch^p which $\to 0$ as $h \to 0$.
- The approximation $f^{(r)}(x)$ contains h^r in the denominator.
- As h is successively decreased to smaller values, the truncation error decreases, but the round-off error in the method may increase as we are dividing by a small number.
- It may happen that after a certain critical value of h, the round-off error may become more dominant than the truncation error and the numerical results obtained may start worsening as h is further reduced.
- when $f(x)$ is given in tabular form, these values may not themselves be exact. These values contain round-off error, i.e.,

$$f(x_i) = f_i + \epsilon_i, \text{ where } f(x_i): \text{ exact value, } f_i: \text{ tabulated value}$$

Choice for h

- To see effect of the round-off error, we consider

$$f'(x_0) = \frac{f(x_1) - f(x_0)}{h} - \frac{h}{2} f''(\zeta), \quad x_0 < \zeta < x_1$$

- If the round off error in $f(x_0)$ and $f(x_1)$ and ϵ_0 and ϵ_1 respectively. Then we have

$$\begin{aligned} f'(x_0) &= \frac{f_1 - f_0}{h} + \frac{\epsilon_1 - \epsilon_0}{h} - \frac{h}{2} f''(\zeta) \\ f'(x_0) &= \frac{f_1 - f_0}{h} + \text{RE(round-off error)} + \text{TE(truncation error)} \end{aligned}$$

- If we take $\epsilon = \max(|\epsilon_1|, |\epsilon_2|)$ and $M_2 = \max\limits_{x_0 \le x \le x_1} |f''(x)|$
- Then we get $|RE| \le \dfrac{2\epsilon}{h}$, $|TE| \le \dfrac{h}{2} M_2$
- Optimal value of h depends on one of the following criteria

$$\text{(i) } |RE| = |TE| \text{ or (ii) } |RE| + |TE| = \text{minimum}$$

Choice for h

- Optimal value of h depends on one of the following criteria

  (i) $|RE| = |TE|$
  (ii) $|RE| + |TE| = \text{minimum}$

- Following first condition, (i), we have $\dfrac{2\epsilon}{h} = \dfrac{h}{2}M_2$
- This gives $h_{adaptive} = 2\sqrt{\epsilon/M_2}$
- Following first condition, (ii), we have

$$\frac{2\epsilon}{h} + \frac{h}{2}M_2 = \text{minimum}, \quad \Rightarrow -\frac{2\epsilon}{h^2} + \frac{1}{2}M_2 = 0 \tag{1}$$

- Thus $h_{adaptive} = 2\sqrt{\epsilon/M_2}$

Example

Example

For the following method, determine the optimal value of h

$$f'(x) = \frac{-3f(x_0) + 4f(x_1) - f(x_2)}{2h} + \frac{h^2}{3}f'''(\zeta), \; x_0 < x < x_2$$

Consider $f(x) = ln(x)$ with tabular value.

x	2.0	2.01	2.02	2.06	2.12
$f(x)$	0.69315	0.69813	0.70310	0.72271	0.75142

with given maximum round-off error in function evaluation is 5×10^{-6}

Section-4.2
Richardson's Extrapolation

Richardson's Extrapolation

- Taylor Series of f about the point x and evaluated at $x+h$ and $x-h$ leads to the central difference formula:
- Solving for f_j' gives: $f_j' = \dfrac{f_{j+1} - f_{i-1}}{2h} + O(h^2)$
- This formula describes precisely how the error behaves.
- This information can be exploited to improve the quality of the numerical solution without ever knowing $f^{(3)}, f^{(5)}, \cdots$
- Let us rewrite this in the following form:

$$f_j' = N_1(h) - \frac{h^2}{6} f_j''' - \frac{h^4}{120} f_j^{(5)} - \cdots$$

where $N_1(h) = \frac{f(x+h) - f(x-h)}{2h}$

- The key of the process is to now replace h by $h/2$ in this formula

Richardson's Extrapolation

- Richardson's extrapolation is used to generate high-accuracy results while using low-order formulas.
  - When an approximation technique has an error term with a predictable form
  - Which depends on a parameter (say, step size h)

Let we have an approximation formula $N_1(h)$ that approximates an unknown constant M

- The truncation error

$$M - N_1(h) = K_1 h + K_2 h^2 + K_3 h^3 + \cdots$$

  where $K_1, K_2, K_3, \ldots$ are unknown constants
- The truncation error is $\mathcal{O}(h)$, i.e., $M - N_1(h) \approx K_1 h$.

Richardson's Extrapolation

- Let $N_2(h)$ is an $\mathcal{O}(h^2)$ approximation formula (obtained by taking a combination of $N_1(h)$ approximations) for M with

$$M - N_2(h) = \hat{K}_2 h^2 + \hat{K}_3 h^3 + \cdots ,$$

  such that $M - N_2(h) \approx \hat{K}_2 h^2$.
- $N_2(h)$ is a better approximation than $N_1(h)$, provided K_1 and $\hat{K}_2$ are comparable in the magnitude.
- Similarly, by taking a combination of $N_2(h)$, we can obtain a third order approximation formula $N_3(h)$, such that, $M - N_3(h) \approx \hat{\hat{K}}_3 h^3$.

Richardson's Extrapolation

How to generate the extrapolation formulas?

- Let $M = N_1(h) + K_1 h + K_2 h^2 + K_3 h^3 + \cdots$
- The above expression is also true for $h/2$, i.e.,

$$M = N_1\left(\frac{h}{2}\right) + K_1\frac{h}{2} + K_2\frac{h^2}{4} + K_3\frac{h^3}{8} + \cdots$$

- Subtracting the first equation with twice of the second equation, we get

$$M = N_1\left(\frac{h}{2}\right) + \left[N_1\left(\frac{h}{2}\right) - N_1(h)\right] + K_2\left(\frac{h^2}{2} - h^2\right) + K_3\left(\frac{h^3}{4} - h^3\right) + \cdots$$

- Hence, we have an $\mathcal{O}(h^2)$ approximation formula for M, i.e.,
$M = N_2(h) - \frac{K_2}{2}h^2 - \frac{3K_3}{4}h^3 - \cdots$, where

$$N_2(h) = N_1\left(\frac{h}{2}\right) + \left[N_1\left(\frac{h}{2}\right) - N_1(h)\right].$$

Example

Use the forward-difference formula to approximate the derivative of $f(x) = \ln x$ at $x_0 = 1.8$ using $h = 0.1$, and $h = 0.05$. Further, use extrapolation on these values

- first-order approximation formula $N_1(h)$ based on forward-difference
- Use $\dfrac{d}{dx}(\ln x) = \dfrac{1}{x}$, we get $f'(1.8) = 1/1.8 = 0.\overline{5}$.
- for $h = 0.1$, $f'(1.8) \approx \dfrac{f(1.9) - f(1.8)}{0.1} = 0.5406722 = N_1(0.1)$ with error 1.5×10^{-2}
- for $h = 0.05$, $\quad f'(1.8) \approx \dfrac{f(1.85) - f(1.8)}{0.05} = 0.5479795 = N_1(0.05)$ with error 7.7×10^{-3}
- Extrapolating these two results to obtain the new approximation

$$N_2(0.1) = N_1(0.05) + (N_1(0.05) - N_1(0.1)) = 0.555287$$

with error 2.7×10^{-4}.

Richardson's Extrapolation

- Extrapolation can be applied whenever the truncation error for a formula has the form $\sum_{j=1}^{m-1} K_j h^{\alpha_j} + \mathcal{O}(h^{\alpha_m})$ for constants K_j and $\alpha_1 < \alpha_2 < \alpha_3 < \cdots < \alpha_m$.
- Some extrapolation formulas have truncation errors with only even powers of h, $M = N_1(h) + K_1 h^2 + K_2 h^4 + \cdots$.
- Subtracting the above equation from four times the equation $M = N_1\left(\frac{h}{2}\right) + K_1 \frac{h^2}{4} + K_2 \frac{h^4}{16} + \cdots$, we get

$$M = \frac{1}{3}\left[4N_1\left(\frac{h}{2}\right) - N_1(h)\right] + \frac{K_2}{3}\left(\frac{h^2}{4} - h^4\right) + \frac{K_3}{3}\left(\frac{h^6}{16} - h^6\right) + \cdots$$

- gives us the fourth-order approximation formula

$$N_2(h) = \frac{1}{3}\left[4N_1\left(\frac{h}{2}\right) - N_1(h)\right] = N_1\left(\frac{h}{2}\right) + \frac{1}{3}\left[N_1\left(\frac{h}{2}\right) - N_1(h)\right]$$

Section-4.5
Romberg Integration

Romberg Integration

- For $f \in C^{\infty}[a, b]$, the composite Trapezoidal rulde can be written with an error term in the form

$$\int_a^b f(x)dx = \frac{h}{2}\left[f(a) + 2\sum_{j=1}^{n-1} f(x_j) + f(b)\right] + K_1 h^2 + K_2 h^4 + K_3 h^6 + \cdots$$

Where constant K_i are depends only on $f^{(2i-1)}(a)$ and $f^{(2i-1)}(b)$

- Here the truncation error has the form

$$\sum_{j=1}^{m-1} K_j h^{2j} + \mathcal{O}(h^{2m}).$$

- Therefore, Richardson extrapolation can be performed.

Romberg Integration

- Use the results of the Composite Trapezoidal rule with $n = 1, 2, 4, 8, 16, \ldots$ to approximate the integral $\int_a^b f(x)dx$, denoted by $R_{1,1}, R_{2,1}, R_{3,1}, \ldots,$ respectively.
- Then obtain $\mathcal{O}(h^4)$ approximations $R_{2,2}, R_{3,2}, R_{4,2}, \ldots,$ where

$$R_{k,2} = R_{k,1} + \frac{1}{3}(R_{k,1} - R_{k-1,1}), \qquad \text{for } k = 2, 3, \ldots$$

- Then the $\mathcal{O}(h^6)$ approximations $R_{3,3}, R_{4,3}, R_{5,3}, \ldots$

$$R_{k,3} = R_{k,2} + \frac{1}{15}(R_{k,2} - R_{k-1,2}), \qquad \text{for } k = 3, 4, \ldots$$

- In general, to obtain the $\mathcal{O}(h^{2j})$ approximations $R_{k,j-1}$

$$R_{k,j} = R_{k,j-1} + \frac{1}{4^{j-1} - 1}(R_{k,j-1} - R_{k-1,j-1}), \qquad \text{for } k = j, j+1, \ldots$$

Example

- To find approximations to $\int_0^\pi \sin x dx$, with $n = 1, 2, 4, 8, 16$

$$R_{1,1} = \frac{\pi}{2}[\sin 0 + \sin \pi] = 0$$
$$R_{2,1} = \frac{\pi}{4}\left[\sin 0 + 2\sin\frac{\pi}{2} + \sin \pi\right] = 1.57079633$$
$$R_{3,1} = \frac{\pi}{8}\left[\sin 0 + 2\left(\sin\frac{\pi}{4} + \sin\frac{\pi}{2} + sin\frac{3\pi}{4}\right) + \sin \pi\right] = 1.8961189$$
$$R_{4,1} = 1.97423160$$
$$R_{5,1} = 1.99357034.$$

- Then using Romberg integration, we obtain the $\mathcal{O}(h^4)$ approximations

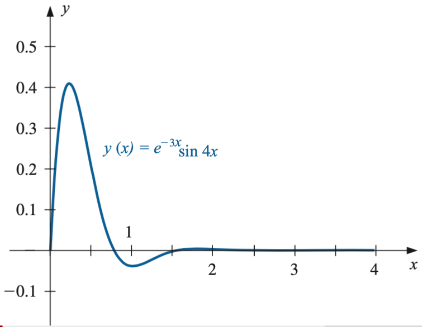
$$R_{2,2} = R_{2,1} + \frac{1}{3}(R_{2,1} - R_{1,1}) = 2.09439511, \quad R_{3,2} = 2.00455976$$
$$R_{4,2} = 2.00026917, \quad R_{5,2} = 2.00001659.$$

Section-4.6
Adaptive Quadrature Methods

Adaptive Quadrature Methods

- Composite formulas suffer occasionally, since they require the use of equally-spaced nodes
- Inappropriate for functions having different regions of large and small functional variations

Example: The solution of the differential equation $y'' + 6y' + 25 = 0$, with $y(0) = 0$ and $y'(0) = 4$ is $y(x) = e^{-3x} \sin 4x$.



Adaptive Quadrature Methods

- To approximate $\int_a^b f(x)dx$, within a tolerance $\varepsilon > 0$
- Use Simpson's rule with step size $h = (b-a)/2$

$$\int_a^b f(x)dx = S(a,b) - \frac{h^5}{90} f^{(4)}(\xi), \quad \text{for some } \xi \in (a,b) \quad (2)$$

where the Simpson's rule approximation is

$$S(a,b) = \frac{h}{3}[f(a) + 4f(a+h) + f(b)]$$

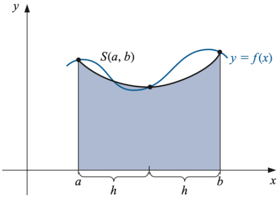
Adaptive Quadrature Methods

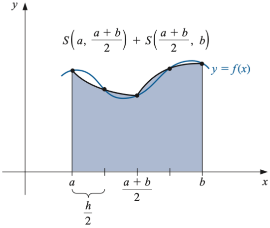
- Next, to determine an accuracy approximation that does not require $f^{(4)}(\xi)$, using Composite Simpson's rule with $N = 4$ and step size $(b-a)/4 = h/2$, giving

$$\int_a^b f(x)dx \approx \frac{h}{6}\left[f(a) + 4f\left(a+\frac{h}{2}\right) + 2f(a+h) + 4f\left(a+\frac{3h}{2}\right) + f(b)\right] \tag{3}$$

$$\approx S\left(a, \frac{a+b}{2}\right) + S\left(\frac{a+b}{2}, b\right)$$

with error $-\frac{1}{16}\left(\frac{h^5}{90}\right) f^{(4)}(\tilde{\xi})$, for some $\tilde{\xi} \in (a, b)$.





Adaptive Quadrature Methods

- From (2) and (3) with $\zeta \approx \tilde{\zeta}$ we get

$$S\left(a, \frac{a+b}{2}\right) - S\left(\frac{a+b}{2}, b\right) - \frac{1}{16}\left(\frac{h^5}{90}\right) f^{(4)}(\xi) \approx S(a, b) - \frac{h^5}{90} f^{(4)}(\xi)$$

This implies

$$\frac{h^5}{90} f^{(4)}(\xi) \approx \frac{16}{15}\left[S(a, b) - S\left(a, \frac{a+b}{2}\right) - S\left(\frac{a+b}{2}, b\right)\right]$$

Adaptive Quadrature Methods

- The above estimate (3) produces the error estimation

$$\left|\int_a^b f(x)dx - S\left(a, \frac{a+b}{2}\right) - S\left(\frac{a+b}{2}, b\right)\right| \approx \frac{1}{16}\left(\frac{h^5}{90}\right) f^{(4)}(\xi)$$
$$\approx \frac{1}{15}\left|S(a,b) - S\left(a, \frac{a+b}{2}\right) - S\left(\frac{a+b}{2}, b\right)\right|$$

- Implies that

$$\left|S\left(a, \frac{a+b}{2}\right) + S\left(\frac{a+b}{2}, b\right) - S(a,b)\right| < 15\varepsilon$$

Therefore, $S(a,(a+b)/2) + S((a+b)/2, b)$ approximates $\int_a^b f(x)dx$ with a sufficient accuracy, such that,

$$\left|\int_a^b f(x)dx - S\left(a, \frac{a+b}{2}\right) - S\left(\frac{a+b}{2}, b\right)\right| < \varepsilon.$$

Example

- For the integral $\int_0^{\pi/2} sinxdx = 1$.
- We have

$$S\left(0, \frac{\pi}{2}\right) = \frac{\pi/4}{3}\left[sin0 + 4sin\frac{\pi}{4} + sin\frac{\pi}{2}\right] = 1.002279878$$

- and

$$S\left(0, \frac{\pi}{4}\right) - S\left(\frac{\pi}{4}, \frac{\pi}{2}\right) = 1.000134585$$

- Therefore,

$$\left|S\left(0, \frac{\pi}{4}\right) + S\left(\frac{\pi}{4}, \frac{\pi}{2}\right) - S\left(0, \frac{\pi}{2}\right)\right| = 0.002145293$$

- Hence, the estimate of the error obtained using $S(a, (a+b)/2) + S((a+b)/2, b)$ to approximate $\int_a^b f(x)dx$ is

$$\frac{1}{15}\left|S\left(0, \frac{\pi}{4}\right) - S\left(\frac{\pi}{4}, \frac{\pi}{2}\right) - S\left(0, \frac{\pi}{2}\right)\right| = 0.000143020.$$


ANY
QUESTIONS?