Revisiting Numerical Integration: Getting More from Fewer Points

Bruce Cohen[1] and David Sklar[2]

[1]Lowell High School
math.cohen@gmail.com
www.cgl.ucsf.edu/home/bic

[2]San Francisco State University
dsklar46@yahoo.com

Asilomar Conference 2007
v0.12 December 30, 2007

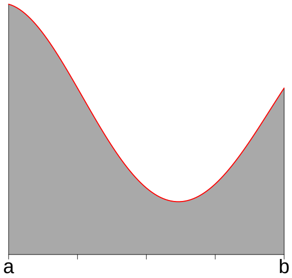
Plan

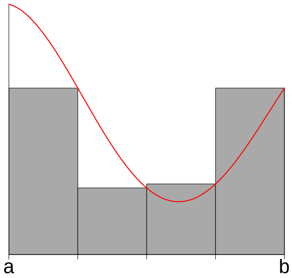
- Worksheet
  - Reviews Midpoint and Trapezoidal methods with attention to the Error
  - u-substitution
  - Why bother with numerical integration
- Simpson's Rule from two perspectives:
  (a) Geometry of errors
  (b) a local second order approximation (hanging a parabola) generalizes to higher order approximations
- A closer look at Midpoint and Trapezoidal errors via Taylor expansion
- Romberg integration: Postprocessing to get more with fewer points
- Gaussian quadrature: Preprocessing to get even more with even fewer points
- Bibliography

What is an Integral?

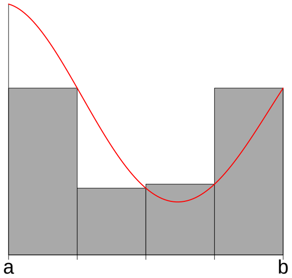
$$\int_a^b f(x)\,dx \qquad = \qquad \lim_{\Delta x \to 0} \sum_{k=1}^{n} f(x_k)\Delta x$$

not $F(b) - F(a)$

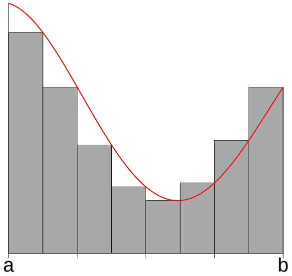


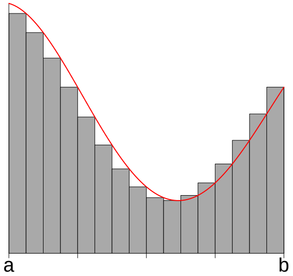


Approximating Integrals

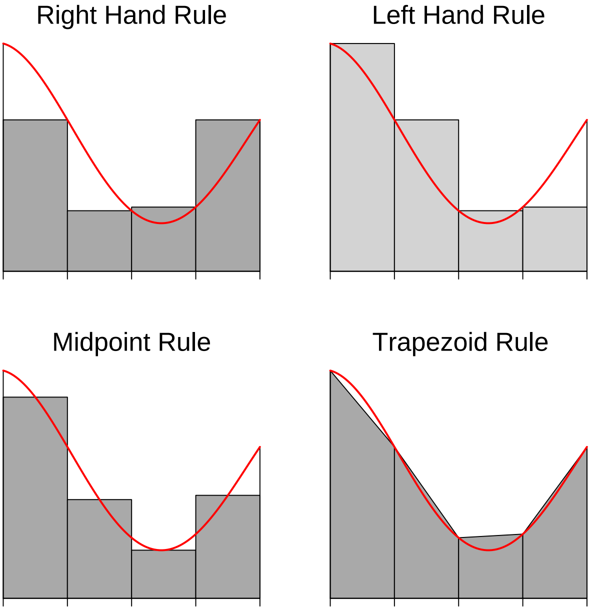


$$\int_a^b f(x)\,dx = \sum_{k=1}^{n} f(x_k)\Delta x + E(\Delta x)$$



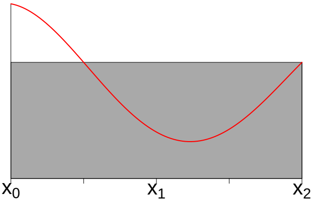


Approximating Area

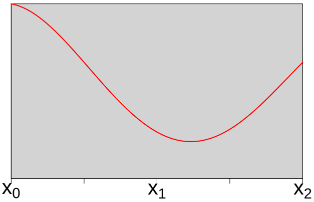


Single Subintervals

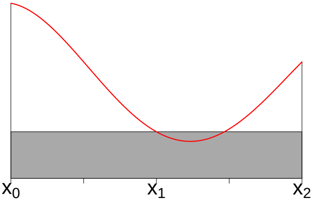
Right Hand Rule



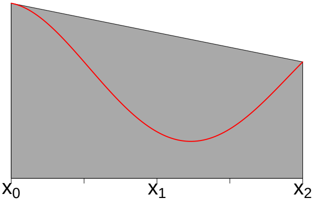
Left Hand Rule



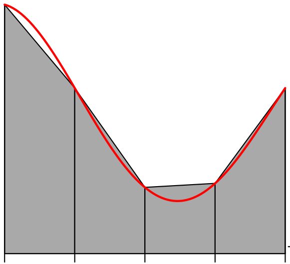
Midpoint Rule



Trapezoid Rule



Quasi Riemann Sums Approximations



$$\int_a^b f(x)\,dx = \sum_{k=1}^{n} w_k f(x_k)\Delta x + E(\Delta x)$$

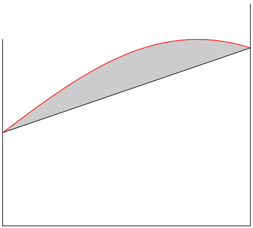
$$\begin{array}{l}
\frac{h}{2}(y_o + y_1) \\
\qquad \frac{h}{2}(y_1 + y_2) \\
\qquad\qquad \frac{h}{2}(y_2 + y_3) \\
\qquad\qquad\qquad \frac{h}{2}(y_3 + y_4) \\
\hline
\frac{h}{2}(y_o + 2\,y_1 + 2\,y_2 + 2\,y_3 + y_4)
\end{array}$$

We associate the weights

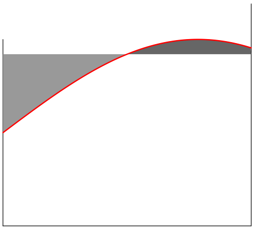
$$(1, 2, 2, \ldots, 2, 1)$$

with the Trapezoid rule.

Geometry of Trapezoid and Midpoint Errors 1

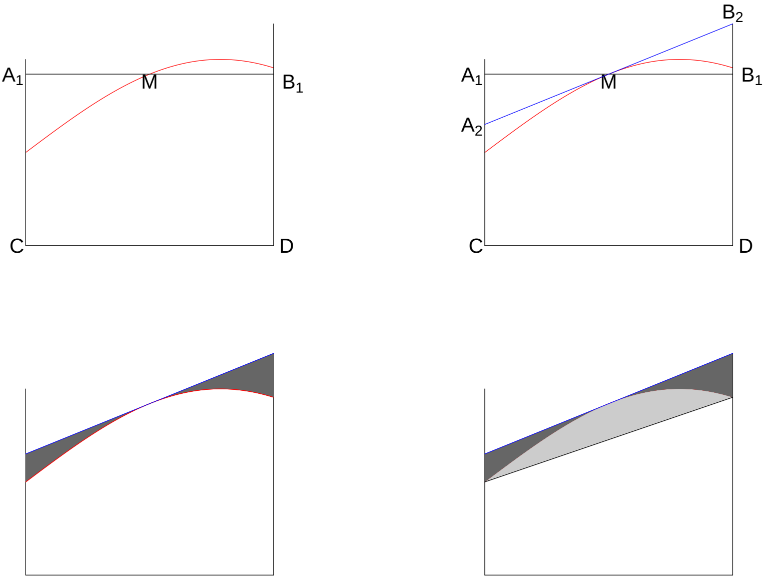


E_T

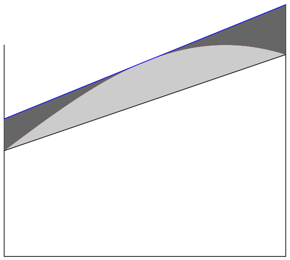


E_M

Geometry of Trapezoid and Midpoint Errors 2



Canceling Errors

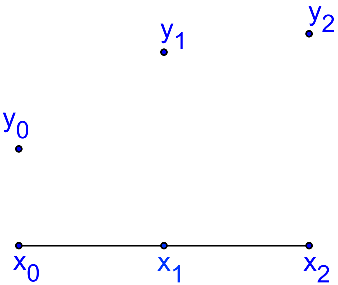


- What can you say about the relative magnitudes of E_T and E_M?
- What can you say about the relative signs of E_T and E_M?

$E_T \approx -2E_M$.

$$\begin{aligned} I &= T + E_T \\ I &= M + E_M \\ I &\approx T - 2E_M \\ 2I &= 2M + 2E_M \\ 3I &\approx 2M + T \\ \therefore I &\approx \frac{2M + T}{3} \end{aligned}$$

Using Simpson's Rule

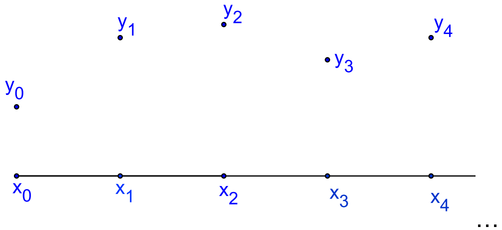


$$S = \frac{2M + T}{3}$$

Assume the subinterval has length $2h$,
$T = h(f(x_0) + f(x_2))$
$M = 2h(f(x_1))$
$S = \frac{1}{3}\left[2 \cdot 2h(f(x_1)) + h(f(x_0) + f(x_2))\right]$
Now pull out the h and rearrange the y values.

$$S = \frac{h}{3}(f(x_0) + 4f(x_1) + f(x_2))$$

Simpson's Rule Weights



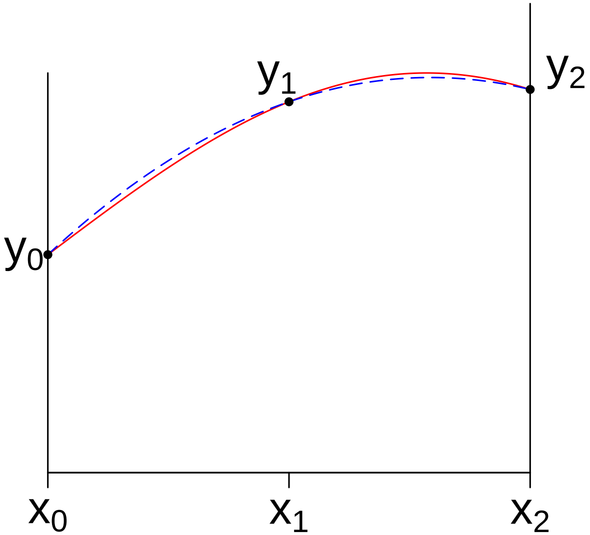
y_0	y_1	y_2	y_3	y_4	y_5	y_6
1	4	1				
		1	4	1		
				1	4	1
1	4	2	4	2	4	1

We associate the weights

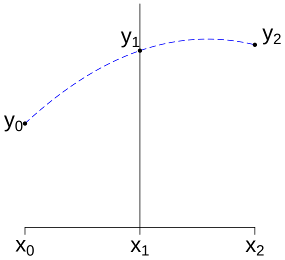
$$(1, 4, 2, 4, 2, \ldots, 4, 1)$$

along with a factor of $\frac{h}{3}$ Simpsons's rule.

Hanging Curves



Building a Parabola for the Area Under



Subinterval $[x_0, x_2]$ has length $2h$ with points (x_0, y_0), (x_1, y_1), and (x_2, y_2). Assume $p(u) = au^2 + bu + c$ is a parabola going through these three points where $u = x_1 - x$.

$c = y_1$

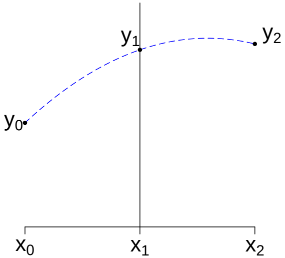
Let A be the area under the parabola.

$$A = \int_{-h}^{h} (au^2 + bu + c)\, du$$

$$= \frac{au^3}{3} + \frac{bu^2}{2} + cu \Big|_{-h}^{h}$$

$$\therefore A = \frac{2ah^3}{3} + 2ch = 2h\left(\frac{ah^2}{3} + c\right)$$

Building a Parabola: Getting a and c



$$A = 2h\left(\frac{ah^2}{3} + c\right)$$

$p(u) = au^2 + bu + c$ is a parabola going through these three points where $u = x_1 - x$.

$$c = y_1$$
$$p(-h) = y_0 = ah^2 - bh + y_1$$
$$p(h) = y_2 = ah^2 + bh + y_1$$
$$y_2 + y_0 = 2ah^2 + 2y_1$$
$$\implies a = \frac{y_0 - 2y_1 + y_2}{2h^2}.$$

Substituting values for a and c:

$$A = 2h\left(\frac{(y_0 - 2y_1 + y_2)h^2}{(2h^2)3} + y_1\right)$$
$$A = 2h\left(\frac{(y_0 - 2y_1 + y_2)}{6} + \frac{6y_1}{6}\right)$$
$$\therefore A = \frac{h}{3}(y_o + 4y_1 + y_2)$$

Simpson's Rule Summary

Notice that our result for A, the area under a three point parabola

$$A = \frac{h}{3}\left(y_o + 4y_1 + y_2\right)$$

matches the equation for S, the weighted midpoint-trapezoid average definition of Simpson's rule.

$$S = \frac{h}{3}\left(f(x_o) + 4f(x_1) + f(x_2)\right)$$

We associate the weights

$$(1, 4, 2, 4, 2, \ldots, 4, 1)$$

along with a factor of $\frac{h}{3}$ Simpsons's rule.

Using Midpoint, Trapezoid, and Simpson's Rule

$$I = \int_1^3 x^{\frac{3}{2}}\,dx$$

$$I \approx 5.835382907248$$

h	$M(h)$	$E_M(h)$	$T(h)$	$E_T(h)$	$S(h)$
$\frac{1}{2}$	5.82397	0.01142	5.85823	-0.02285	5.835389276
$\frac{1}{4}$	5.83252	0.00286	5.84110	-0.00572	5.835383315
$\frac{1}{8}$	5.83467	0.00071	5.83681	-0.00143	5.835382933
$\frac{1}{16}$	5.83520	0.00018	5.83574	-0.00036	5.835382909

Error Analysis Road Map

1. Obtain the error on a single subinterval in terms of h:
   a. expand $f(x)$ in a Taylor series about the midpoint
   b. express the integral in terms of h using the series
   c. express the function values at the endpoints in terms of h using the series
   d. express the numerical approximation in terms of h using the previous results
   e. express the error (integral minus approximation) in terms of h
2. Obtain the total error for the entire integral in terms of h:
   a. add the errors for the subintervals
   b. argue that certain key expressions that involve h are actually independent of h
   c. determine the coefficient of the leading term of the series for the error

Midpoint Error on One Subinterval

- a single subinterval $[x_0, x_2]$ with midpoint x_1
- $x_2 - x_0 = h$
- Our task is to look for the difference between the exact value of the integral $I = \int_{x_0}^{x_2} f(x)\,dx$ and an approximation on one subinterval.
- $f(x)$ is sufficiently nice that the Taylor series centered at x_1 converges everywhere on $[x_0, x_2]$.

$$f(x) = \sum_n \frac{f^{(n)}(x_1)}{n!}(x - x_1)^n$$

By letting $u = x - x_1$ the graph of f over $[x_o, x_2]$ is the same as the graph of $g(u) = \sum_n \frac{f^{(n)}(x_1)}{n!} u^n$ for $u \in [-h/2, h/2]$.

Midpoint Error on One Subinterval: Getting I

$$I = \int_{x_0}^{x_2} f(x)\,dx = \int_{-h/2}^{h/2} g(u)\,du$$

$$I = \int_{-h/2}^{h/2} \Big[\sum_n \frac{f^{(n)}(x_1)}{n!} u^n\,du \Big] = \sum_n \Big[\frac{f^{(n)}(x_1)}{n!} \int_{-h/2}^{h/2} u^n\,du \Big]$$

Notice that the symmetry of an odd function u^n over the interval $[-h/2, h/2]$ eliminates the odd terms of the expansion of I.

Algebraically, $\displaystyle\int_{-h/2}^{h/2} u^n\,du = \frac{u^{n+1}}{n+1}\Big|_{-h/2}^{h/2} = \frac{(h/2)^{n+1} - (-h/2)^{n+1}}{n+1}$

so the terms sum to zero when $n+1$ is even (i.e. when n is odd).

$$I = \sum_{n \textbf{ even}} \frac{f^{(n)}(x_1)}{n!} \frac{2h^{n+1}}{(n+1)2^{n+1}}$$

$$I = h \sum_{n \textbf{ even}} \frac{f^{(n)}(x_1)}{2^n (n+1)!} h^n$$

Midpoint Error on One Subinterval: Getting $E_{M_1}(h)$

Take a look at the first term of this expansion:

$$I = h \cdot f(x_1) + h \sum_{\substack{n \textbf{ even} \\ n>0}} \frac{f^{(n)}(x_1)}{2^n(n+1)!} h^n$$

Notice $h \cdot f(x_1)$ is $M(h)$, the one subinterval approximation for I

$$E_{M_1}(h) = I - M(h) = h \sum_{\substack{n \textbf{ even} \\ n>0}} \frac{f^{(n)}(x_1)}{2^n(n+1)!} h^n$$

$$E_{M_1}(h) = h \left[\frac{1}{24} f''(x_1) h^2 + b_4 h^4 + b_6 h^6 + ... \right]$$

$$\therefore E_{M_1}(h) = h \left[\frac{1}{24} f''(x_1) h^2 + \mathcal{O}(h^4) \right]$$

Trapezoid Error on One Subinterval: $T(h)$ from Taylor Expansion

The trapezoid approximation for one subinterval is:

$$T(h) = \frac{h}{2}[f(x_o) + f(x_2)]$$

We can use the Taylor expansion of $f(x)$ to rewrite the trapezoid approximation.

$$f(x_o) = \sum_n \frac{f^{(n)}(x_1)}{n!}\left(-\frac{h}{2}\right)^n$$

$$f(x_2) = \sum_n \frac{f^{(n)}(x_1)}{n!}\left(\frac{h}{2}\right)^n$$

$$\frac{1}{2}(f(x_0) + f(x_2)) = \frac{1}{2}\sum_{n \textbf{ even}} \frac{f^{(n)}(x_1)}{n!}\frac{h^n}{2^{n-1}}$$

$$\textbf{so } T(h) = h\left(\sum_{n \textbf{ even}} \frac{f^{(n)}(x_1)}{n!}\frac{h^n}{2^n}\right)$$

Trapezoid Error on One Subinterval: $E_{T_1}(h)$

Noting that $I = T(h) + E_T(h)$ (or $E_T(h) = I - T(h)$) we use previous equations:

$$E_{T_1}(h) = h\left(\sum_{n \textbf{ even}} \frac{f^{(n)}(x_1)}{2^n(n+1)!}h^n - \sum_{n \textbf{ even}} \frac{f^{(n)}(x_1)}{n!}\frac{h^n}{2^n}\right)$$

$$= h\left(\sum_{n \textbf{ even}} \frac{-n}{2^n(n+1)!} f^{(n)}(x_1)h^n\right)$$

$$\therefore E_{T_1}(h) = h\left[\frac{-1}{12} f''(x_1)h^2 + \mathcal{O}(h^4)\right].$$

Total Midpoint Error

In order to get the total midpoint error, we need to sum the error for each subinterval. Staying with our x_k notation, assume n subintervals with $2n+1$ equally spaced x-values where $x_0 = a$, $x_{2n} = b$ and any x_{2k+1} ("odd subscripted" x-value) is in the middle of a subinterval. Generalizing the single subinterval error equation gives:

$$E_{M_j}(h) = h\left[\frac{1}{24}f''(x_{2j-1})h^2 + \mathcal{O}(h^4)\right]$$

So adding these single subinterval errors gives:

$$E_M(h) = \sum_{j=1}^{n} E_{M_j}(h)$$

$$E_M(h) = \sum_{j=1}^{n} h\left[\frac{1}{24}f''(x_{2j-1})h^2 + \mathcal{O}(h^4)\right]$$

$$E_M(h) = \frac{h^2}{24}\sum_{j=1}^{n} f''(x_{2j-1})h + \left(\sum_{j=1}^{n} h\right)\mathcal{O}(h^4)$$

Total Midpoint Error: No Worse than $\mathcal{O}(h^2)$

$$E_M(h) = \frac{h^2}{24}\sum_{j=1}^{n} f''(x_{2j-1})h + \left(\sum_{j=1}^{n} h\right)\mathcal{O}(h^4)$$

Since h is assumed to be a small number strictly between zero and one, we can safely say that $E_M(h)$ will be no worse than $\mathcal{O}(h^2)$. Consider the second term:

$$\sum_{j=1}^{n} h = b - a.$$

This leaves the second term as $(b-a)$, a constant, times a series whose largest term is of order 4.

$$E_M(h) = \frac{h^2}{24}\sum_{j=1}^{n} f''(x_{2j-1})h + \mathcal{O}(h^4)$$

Total Midpoint Error: Bootstrap and FTC

Now notice that the sum in the first term $\left(\sum_{j=1}^{n} f''(x_{2j-1})h\right)$ is just a midpoint approximation for

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

Since the midpoint error is no worse than $\mathcal{O}(h^2)$:

$$\sum_{j=1}^{n} f''(x_{2j-1})h = f'(b) - f'(a) + E_M(h) = f'(b) - f'(a) + \mathcal{O}(h^2),$$

so,

$$E_M(h) = \frac{h^2}{24}\left(f'(b) - f'(a) + \mathcal{O}(h^2)\right) + \mathcal{O}(h^4)$$
$$\therefore E_M(h) = \frac{f'(b) - f'(a)}{24} h^2 + \mathcal{O}(h^4)$$

Total Trapezoid Error

Following the path just detailed for the total midpoint error, the path to a total trapezoid error starts with a generalized subinterval error:

$$E_{T_j}(h) = h\left[\frac{-1}{12}f''(x_{2j-1})h^2 + \mathcal{O}(h^4)\right]$$

and ends with

$$E_T(h) = -\frac{f'(b) - f'(a)}{12}h^2 + \mathcal{O}(h^4)$$

Total Midpoint and Trapezoid Errors

$$E_M(h) = \frac{f'(b) - f'(a)}{24} h^2 + b_4 h^4 + b_6 h^6 + \ldots$$

$$E_T(h) = -\frac{f'(b) - f'(a)}{12} h^2 + c_4 h^4 + c_6 h^6 + \ldots$$

Notice that indeed the trapezoid error is exactly twice the magnitude of the midpoint error (at least until the order of h^4) and opposite in sign.