Simple Numerical Integrators – Error Behaviour

These notes provide an introduction to the error behaviour of the Midpoint, Trapezoidal and Simpson's Rules for generating approximate values of the definite integral $\int_a^b f(x)\,dx$. These rules are, in order

$$\int_a^b f(x)\,dx \approx \Big[f(\bar{x}_1) + f(\bar{x}_2) + \cdots + f(\bar{x}_n)\Big]\Delta x \tag{M}$$

$$\int_a^b f(x)\,dx \approx \Big[\tfrac{1}{2}f(x_0) + f(x_1) + f(x_2) + \cdots + f(x_{n-1}) + \tfrac{1}{2}f(x_n)\Big]\Delta x \tag{T}$$

$$\begin{aligned}\int_a^b f(x)\,dx \approx \Big[f(x_0) &+ 4f(x_1) + 2f(x_2) + 4f(x_3) + 2f(x_4) + \cdots \\ &+ 2f(x_{n-2}) + 4f(x_{n-1}) + f(x_n)\Big]\tfrac{\Delta x}{3}\end{aligned} \tag{S}$$

where

$$\Delta x = \tfrac{b-a}{n}, \quad x_0 = a, \quad x_1 = a + \Delta x, \quad x_2 = a + 2\Delta x, \quad \cdots, \quad x_{n-1} = b - \Delta x, \quad x_n = b$$
$$\bar{x}_1 = \tfrac{x_0+x_1}{2}, \quad \bar{x}_2 = \tfrac{x_1+x_2}{2}, \quad \cdots, \quad \bar{x}_n = \tfrac{x_{n-1}+x_n}{2}$$

Two obvious considerations in deciding whether or not a given algorithm is of any practical value are (a) the amount of computational effort required to execute the algorithm and (b) the accuracy that this computational effort yields. For algorithms like our simple integrators, the bulk of the computational effort usually goes into evaluating the function $f(x)$. The number of evaluations of $f(x)$ required for n steps of the Midpoint Rule is n, while the number required for n steps of the Trapezoidal and Simpson's Rules is $n+1$. So all three of our rules require essentially the same amount of effort – one evaluation of $f(x)$ per step.

To get a first impression of the error behaviour of these methods, we apply them to a problem that we know the answer to. The exact value of the integral $\int_0^\pi \sin x\,dx = -\cos x\big|_0^\pi$ is 2. The following table lists the error in the approximate value for this number generated by our three rules applied with three different choices of n. It also lists the number of evaluations of f required to compute the approximation.

	Midpoint		**Trapezoidal**		**Simpson's**	
n	error	#evals	error	#evals	error	#evals
10	4.1×10^{-1}	10	8.2×10^{-1}	11	5.5×10^{-3}	11
100	4.1×10^{-3}	100	8.2×10^{-3}	101	5.4×10^{-7}	101
1000	4.1×10^{-5}	1000	8.2×10^{-5}	1001	5.5×10^{-11}	1001

Observe that

- Using 101 evaluations of f worth of Simpson's rule gives an error 80 times smaller than 1000 evaluations of f worth of the Midpoint Rule.

- The Trapezoidal Rule error with n steps is about twice the Midpoint Rule error with n steps.
- With the Midpoint Rule, increasing the number of steps by a factor of 10 appears to reduce the error by about a factor of $100 = 10^2 = n^2$.
- With the Trapezoidal Rule, increasing the number of steps by a factor of 10 appears to reduce the error by about a factor of $10^2 = n^2$.
- With Simpson's Rule, increasing the number of steps by a factor of 10 appears to reduce the error by about a factor of $10^4 = n^4$.

So it looks like

$$\boxed{\begin{aligned}\text{approx value of } \textstyle\int_a^b f(x)\,dx \text{ given by } n \text{ Midpoint steps } &\approx \textstyle\int_a^b f(x)\,dx + K_M \frac{1}{n^2}\\ \text{approx value of } \textstyle\int_a^b f(x)\,dx \text{ given by } n \text{ Trapezoidal steps } &\approx \textstyle\int_a^b f(x)\,dx + K_T \frac{1}{n^2}\\ \text{approx value of } \textstyle\int_a^b f(x)\,dx \text{ given by } n \text{ Simpson's steps } &\approx \textstyle\int_a^b f(x)\,dx + K_M \frac{1}{n^4}\end{aligned}}$$

with some constants K_M, K_T and K_S. It also looks like $K_T \approx 2K_M$.

To test these conjectures further, I have applied our three rules with about ten different choices of n of the form $n = 2^m$ with m integer. On the next page are two figures, one containing the results for the Midpoint and Trapezoidal Rules and the other the results for Simpson's Rule. For each of these rules we are expecting the error e_n (that is, $|\,\text{exact value} - \text{approximate value}\,|$) with n steps to be of the form

$$e_n = K\tfrac{1}{n^k}$$

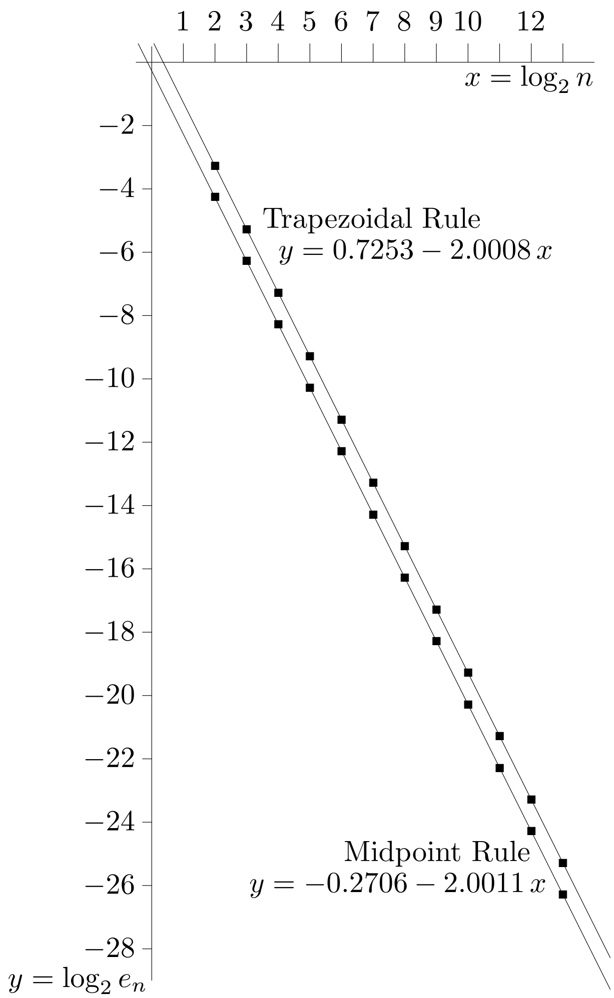
We would like to test if this is really the case. It is not easy to tell whether or not a given curve really is a parabola $y = x^2$ or a quartic $y = x^4$. But the eye is pretty good at determining whether or not a graph is a straight line. Fortunately, there is a little trick that turns the curve $e_n = K\frac{1}{n^k}$ into a straight line – no matter what k is. Instead of plotting e_n against n, plot $\log e_n$ against $\log n$. If $e_n = K\frac{1}{n^k}$, then $\log e_n = \log K - k \log n$. So plotting $y = \log e_n$ against $x = \log n$ gives the straight line $y = \log K - kx$, which has slope $-k$ and y–intercept $\log K$.[†]

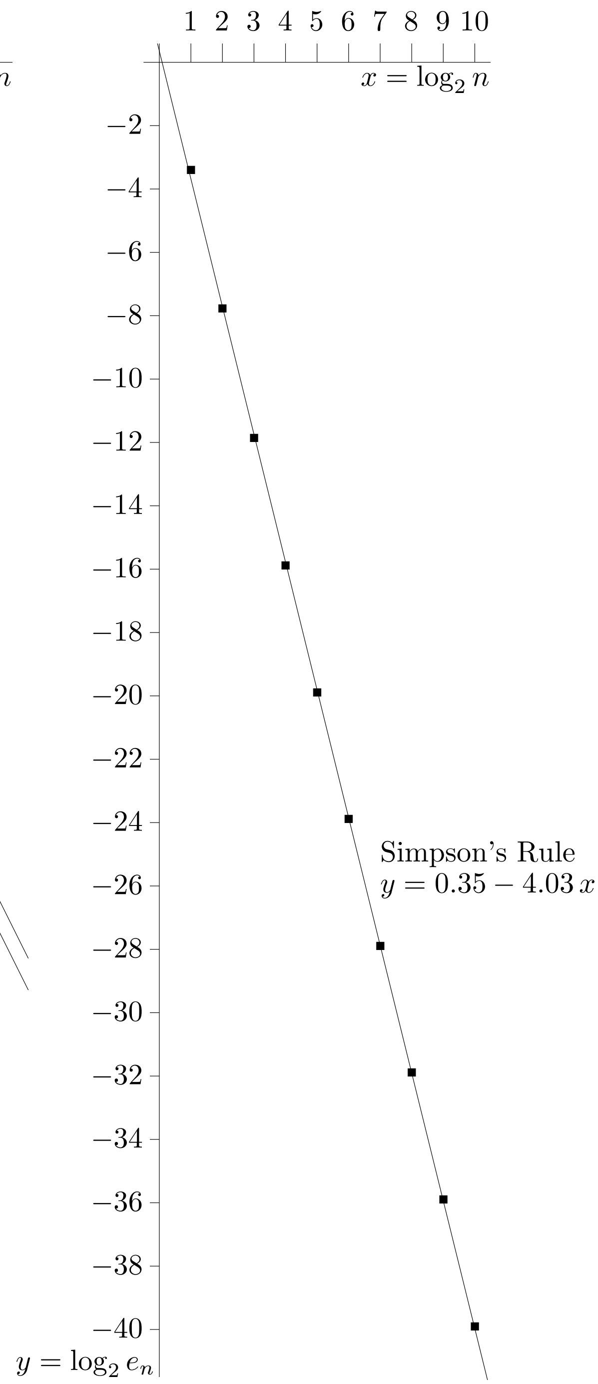
The three graphs on the next page plot $y = \log_2 e_n$ against $x = \log_2 n$ for our three rules. I have chosen to use the base 2 logarithm only because $\log_2 n = \log_2 2^m = m$ is nice and simple. For example, applying Simpson's Rule with $n = 2^5 = 32$ gives the approximate value 2.00000103, which has error $e_n = 0.00000103$. So, I included the data point $(x = \log_2 2^5 = 5, y = \log_2 0.00000103 = \frac{\ln 0.00000103}{\ln 2} = -19.9)$ on the Simpson's Rule graph. For each of the three sets of data points, I have also plotted a straight line "through" the data points. I used linear regression to decide precisely which straight line to plot. Linear regression is not part of this course. It provides a formula for the slope and y–intercept of the straight line which best fits any given set of data points. From the three

[†] There is a variant of this trick that works even when you don't know the answer to the integral ahead of time. Suppose that you suspect that $M_n = A + K\frac{1}{n^k}$, where A is the exact value of the integral and suppose that you don't know the values of A, K and k. Then $M_n - M_{2n} = K\frac{1}{n^k} - K\frac{1}{(2n)^k} = K\big(1 - \frac{1}{2^k}\big)\frac{1}{n^k}$, so plotting $y = \log(M_n - M_{2n})$ against $x = \log n$ gives the straight line $y = \log K\big(1 - \frac{1}{2^k}\big) - kx$.

lines, it sure looks like $k = 2$ for the Midpoint and Trapezoidal Rules and $k = 4$ for Simpson's Rule. It also looks like ratio between the value of K for the Trapezoidal Rule, namely $K = 2^{0.7253}$, and the value of K for the Midpoint Rule, namely $K = 2^{-0.2706}$, is pretty close to 2: $2^{0.7253}/2^{-0.2706} = 2^{0.9959}$.

Error in the approximation, with n steps, to $\int_0^\pi \sin x\,dx$





An Error Bound for the Midpoint Rule

We now try develop some understanding as to why we got the above experimental results. We start with the error generated by a single step of the Midpoint Rule. That is, the error introduced by the approximation

$$\int_{x_0}^{x_1} f(x)\,dx \approx f(\bar{x}_1)\Delta x \qquad \text{where } \Delta x = x_1 - x_0,\ \bar{x}_1 = \tfrac{x_0+x_1}{2}$$

We can get a pretty good idea as to how big this error is by using the quadratic approximation to $f(x)$ about $\bar{x}_1$:

$$f(x) \approx f(\bar{x}_1) + f'(\bar{x}_1)(x-\bar{x}_1) + \tfrac{1}{2}f''(\bar{x}_1)(x-\bar{x}_1)^2$$

In fact we can do even better than this. There is an exact formula for $f(x)$ that looks a lot like the quadratic approximation.

$$f(x) = f(\bar{x}_1) + f'(\bar{x}_1)(x-\bar{x}_1) + \tfrac{1}{2}f''\big(\xi(x)\big)(x-\bar{x}_1)^2$$

Here $\xi(x)$ is some (unknown) number that depends on x and is somewhere between $\bar{x}_1$ and x. Subbing this exact formula into the integral gives

$$\begin{aligned}
\int_{x_0}^{x_1} f(x)\,dx &= f(\bar{x}_1)\int_{x_0}^{x_1} 1\,dx + f'(\bar{x}_1)\int_{x_0}^{x_1}(x-\bar{x}_1)\,dx + \tfrac{1}{2}\int_{x_0}^{x_1} f''\big(\xi(x)\big)(x-\bar{x}_1)^2\,dx \\
&= f(\bar{x}_1)\ x\Big|_{x_0}^{x_1} + f'(\bar{x}_1)\ \tfrac{1}{2}(x-\bar{x}_1)^2\Big|_{x_0}^{x_1} + \tfrac{1}{2}\int_{x_0}^{x_1} f''\big(\xi(x)\big)(x-\bar{x}_1)^2\,dx \\
&= f(\bar{x}_1)\Delta x + \tfrac{1}{2}\int_{x_0}^{x_1} f''\big(\xi(x)\big)(x-\bar{x}_1)^2\,dx
\end{aligned}$$

since $x_1 - x_0 = \Delta x$ and $(x_1-\bar{x}_1)^2 = (x_0-\bar{x}_2)^2 = \big(\frac{\Delta x}{2}\big)^2$. Note that the first term on the right hand side is exactly the Midpoint Rule approximation. So the error that the Midpoint Rule introduces in the first slice is exactly

$$\Big|\int_{x_0}^{x_1} f(x)\,dx - f(\bar{x}_1)\Delta x\Big| = \tfrac{1}{2}\Big|\int_{x_0}^{x_1} f''\big(\xi(x)\big)(x-\bar{x}_1)^2\,dx\Big|$$

Even though we don't know exactly what $\xi(x)$ is, this is a very useful formula if we know that $|f''(x)|$ is smaller than some number M for all x in the domain of integration. For example, when $f(x)=\sin x$, as in the example that we have been following, we know that $|f''(x)| = |-\sin x| \le M = 1$ for all x. Then, $|f''\big(\xi(x)\big)| \le M$, no matter what $\xi(x)$ is, and

$$\begin{aligned}
\Big|\int_{x_0}^{x_1} f(x)\,dx - f(\bar{x}_1)\Delta x\Big| &\le \tfrac{1}{2}\int_{x_0}^{x_1} M(x-\bar{x}_1)^2\,dx \le \tfrac{1}{2}M\ \tfrac{1}{3}(x-\bar{x}_1)^3\Big|_{x_0}^{x_1} \\
&= \tfrac{M}{6}\Big[\big(x_1-\bar{x}_1\big)^3 - \big(x_0-\bar{x}_1\big)^3\Big] = \tfrac{M}{6}\Big[\big(\tfrac{\Delta x}{2}\big)^3 - \big(-\tfrac{\Delta x}{2}\big)^3\Big] \\
&= \tfrac{M}{24}\Delta x^3
\end{aligned}$$

This is a bound on the error introduced by the Midpoint Rule in a single step. When there are n steps, $\Delta x = \frac{b-a}{n}$ so that the error introduced in a single step is bounded by $\frac{M}{24}\big(\frac{b-a}{n}\big)^3$ and the total error is bounded by

$$\left|\int_a^b f(x)\,dx - [f(\bar{x}_1) + \cdots + f(\bar{x}_n)]\Delta x\right| \le n\tfrac{M}{24}\big(\tfrac{b-a}{n}\big)^3 = \boxed{\tfrac{M}{24}\tfrac{(b-a)^3}{n^2}}$$

We have just shown that, if $|f''(x)| \le M$ for all x in the domain of integration,

$$\boxed{\text{the total error introduced by the Midpoint Rule is bounded by } \tfrac{M}{24}\tfrac{(b-a)^3}{n^2}}$$

A similar analysis shows that

$$\boxed{\text{the total error introduced by the Trapezoidal Rule is bounded by } \tfrac{M}{12}\tfrac{(b-a)^3}{n^2}}$$

and if $|f^{(4)}(x)| \le M$ for all x in the domain of integration,

$$\boxed{\text{the total error introduced by Simpson's Rule is bounded by } \tfrac{M}{180}\tfrac{(b-a)^5}{n^4}}$$

For the example $\int_0^\pi \sin x\,dx$, $b-a=\pi$ and M, the largest possible value of $\left|\frac{d^2}{dx^2}\sin x\right|$ (for the Midpoint and Trapezoidal Rules) or $\left|\frac{d^4}{dx^4}\sin x\right|$ (for Simpson's Rule) is 1. So, for the Midpoint Rule,

$$|e_n| \le \tfrac{M}{24}\tfrac{(b-a)^3}{n^2} = \tfrac{\pi^3}{24}\tfrac{1}{n^2} = 1.29\tfrac{1}{n^2}$$

The data in the graph on page 3 gives $|e_n| \approx 2^{-.2706}\frac{1}{n^2} = 0.83\frac{1}{n^2}$ which is consistent with the bound $|e_n| \le 1.29\frac{1}{n^2}$.

In a typical application, one is required to evaluate a given integral to some specified accuracy. For example, if you are manufacturer and your machinery can only cut materials to an accuracy of $\frac{1}{10}^{\rm th}$ of a millimeter, there is no point in making design specifications more accurate than $\frac{1}{10}^{\rm th}$ of a millimeter. Suppose, for example, that we wish to use the Midpoint Rule to evaluate $\int_0^1 e^{-x^2}\,dx$ to within an accuracy of 10^{-6}. (In fact this integral cannot be evaluated exactly, so one must use numerical methods.) The first two derivatives of the integrand are

$$\tfrac{d}{dx}e^{-x^2} = -2xe^{-x^2} \qquad \text{and} \qquad \tfrac{d^2}{dx^2}e^{-x^2} = \tfrac{d}{dx}\big(-2xe^{-x^2}\big) = -2e^{-x^2} + 4x^2e^{-x^2} = 2(2x^2-1)e^{-x^2}$$

As x runs from 0 to 1, $2x^2-1$ runs from -1 to 1, so that

$$0 \le x \le 1 \implies |2x^2-1| \le 1,\ e^{-x^2} \le 1 \implies \left|2(2x^2-1)e^{-x^2}\right| \le 2$$

So the error introduced by the n step Midpoint Rule is at most $\frac{M}{24}\frac{(b-a)^3}{n^2} \le \frac{2}{24}\frac{(1-0)^3}{n^2} = \frac{1}{12n^2}$. This error is at most 10^{-6} if

$$\tfrac{1}{12n^2} \le 10^{-6} \iff n^2 \ge \tfrac{1}{12}10^6 \iff n \ge \sqrt{\tfrac{1}{12}10^6} = 288.7$$

So 289 steps of the Midpoint Rule will do the job.

Suppose now that we wish to use Simpson's Rule to evaluate $\int_0^1 e^{-x^2}\,dx$ to within an accuracy of 10^{-6}. To determine the number of steps required, we must determine how big $\frac{d^4}{dx^4}e^{-x^2}$ can get when $0 \le x \le 1$.

$$\begin{aligned}
\tfrac{d^3}{dx^3}e^{-x^2} &= \tfrac{d}{dx}\big(2(2x^2-1)e^{-x^2}\big) = 8xe^{-x^2} - 4x(2x^2-1)e^{-x^2} = 4(-2x^3+3x)e^{-x^2}\\
\tfrac{d^4}{dx^4}e^{-x^2} &= \tfrac{d}{dx}\big(4(-2x^3+3x)e^{-x^2}\big) = 4(-6x^2+3)e^{-x^2} - 8x(-2x^3+3x)e^{-x^2} = 4(4x^4-12x^2+3)e^{-x^2}
\end{aligned}$$

We now have to find an M such that $g(x) = 4(4x^4-12x^2+3)e^{-x^2}$ obeys $|g(x)| \le M$ for all $0 \le x \le 1$. Here are three different methods for finding such an M.

Method 1: The first method is to find the largest and small value that $g(x)$ takes on the interval $0 \le x \le 1$ by checking the values of $g(x)$ at its critical points and at the end points of the interval of interest. I warn you that, while this method gives the smallest possible value of M, it involves a lot more work than the other methods. It is **not recommended**. Since

$$g'(x) = 4(16x^3-24x)e^{-x^2} - 8x(4x^4-12x^2+3)e^{-x^2} = -8x(4x^4-20x^2+15)e^{-x^2}$$

the critical points of $g(x)$ are $x = 0$ and

$$x^2 = \tfrac{20\pm\sqrt{400-4\times4\times15}}{8} = \tfrac{20\pm\sqrt{160}}{8} = \tfrac{5\pm\sqrt{10}}{2} = 4.081139, 0.918861 \Longrightarrow x = \pm2.020183, \pm0.958572$$

Since

$$g(0) = 12,\ g(0.958572) = -7.419481,\ g(1) = -20e^{-1} = -7.357589$$

we know that $g(x)$ only takes values between -7.419481 and 12, so we may choose $M = 12$.

Method 2: Consider the three factors 4, $4x^4-12x^2+3$, and e^{-x^2} of $g(x)$ separately. For $0 \le x \le 1$, $e^{-x^2} \le e^{-0} = 1$ and

$$\big|4x^4-12x^2+3\big| \le 4x^4+12x^2+3 \le 4+12+3 = 19$$

Hence

$$0 \le x \le 1 \Longrightarrow |g(x)| \le 4\big|4x^4-12x^2+3\big|e^{-x^2} \le 4\times19\times1 = 76$$

So we may choose $M = 76$.

Method 3: Again consider the three factors 4, $4x^4-12x^2+3$ and e^{-x^2} of $g(x)$ separately. But this time, consider the positive terms of $4x^4-12x^2+3$ and the negative terms of $4x^4-12x^2+3$ separately. For $0 \le x \le 1$,

$$3 \le 4x^4+3 \le 7 \text{ and } -12 \le -12x^2 \le 0$$

Adding these two inequalities together gives

$$-9 \le 4x^4-12x^2+3 \le 7$$

Consequently, the maximum value of $|4x^4-12x^2+3|$ for $0 \le x \le 1$ is no more than 9 and

$$\big|g(x)\big| \le 4\times9\times1 = 36$$

We have now found three different possible values of M – all are allowed. In general, the error introduced by the n step Simpson's Rule is at most $\frac{M}{180}\frac{(b-a)^5}{n^4}$. In this example, $a = 0$ and $b = 1$ so that this error is at most 10^{-6} if

$$\frac{M}{180n^4} \le 10^{-6} \iff n^4 \ge \frac{M}{180}10^6 \iff n \ge \sqrt[4]{\frac{M}{180}10^6} = \begin{cases} 16.1 & \text{if } M = 12 \\ 21.1 & \text{if } M = 36 \\ 25.5 & \text{if } M = 76 \end{cases}$$

So if we take $M = 12$, we conclude that 18 steps of the Simpson's Rule will do the job. If we take $M = 36$, we conclude that 22 steps will do the job and if we take $M = 76$, we conclude that 26 steps will do the job. This is a typical case. Method 1 gives a slightly smaller of n than the much simpler procedures of Methods 2 and 3. But usually this gain in n is not worth the extra effort required to apply Method 1.