TRAPEZOIDAL METHOD: ERROR FORMULA

Theorem Assume $f(x)$ twice continuously differentiable on the interval $[a, b]$. Then

$$E_n^T(f) := \int_a^b f(x)\,dx - T_n(f) = -\frac{h^2(b-a)}{12} f''(c_n)$$

for some c_n in the interval $[a, b]$.

Later we will say something about the proof of this result, as it leads to some other useful formulas for the error.

The above formula says that the error decreases in a manner that is roughly proportional to h^2. Thus doubling n (and halving h) should cause the error to decrease by a factor of approximately 4. This is what we observed with some past examples from the preceding section.

Example

Consider evaluating

$$I = \int_0^2 \frac{dx}{1+x^2}$$

using the trapezoidal method $T_n(f)$. Let us bound the error

$$E_n^T(f) = -\frac{h^2(b-a)}{12} f''(c_n)$$

Here, $b - a = 2$. We bound $|f''(c_n)|$ by $\max_{0 \le x \le 2} |f''(x)|$.
Calculate the derivatives:

$$f'(x) = \frac{-2x}{(1+x^2)^2}, \quad f''(x) = \frac{-2+6x^2}{(1+x^2)^3}, \quad f'''(x) = \frac{24x(1-x^2)}{(1+x^2)^4}$$

For $x \in (0, 2)$, $f'''(x) = 0$ only when $x = 1$. So

$$\max_{0 \le x \le 2} |f''(x)| = \max\{|f''(0)|, |f''(1)|, |f''(2)|\} = 2$$

Therefore,

$$\left|E_n^T(f)\right| \le \frac{h^2(2)}{12} \cdot 2 = \frac{h^2}{3}$$

$$\left|E_n^T(f)\right| \le \frac{h^2}{3}$$

How large should n be chosen in order to ensure that

$$\left|E_n^T(f)\right| \le 5 \times 10^{-6} \tag{1}$$

To ensure this, we choose h so small that

$$\frac{h^2}{3} \le 5 \times 10^{-6}$$

This is equivalent to choosing h and n to satisfy

$$h \le .003873$$

$$n = \frac{2}{h} \ge 516.4$$

Thus $n \ge 517$ will imply (1).

DERIVING THE ERROR FORMULA

There are two stages in deriving the error:
(1) Obtain the error formula for the case of a single subinterval $(n = 1)$;
(2) Use this to obtain the general error formula given earlier.

For the trapezoidal method with only a single subinterval, we have

$$\int_{\alpha}^{\alpha+h} f(x)\,dx - \frac{h}{2}\left[f(\alpha) + f(\alpha + h)\right] = -\frac{h^3}{12} f''(c)$$

for some c in the interval $[\alpha, \alpha + h]$.

This error formula can be derived through an application of Taylor's theorem or the Newton form of the linear interpolation error.

The general trapezoidal rule $T_n(f)$ was obtained by applying the simple trapezoidal rule to a subdivision of the original interval of integration. Recall the notation

$$h = \frac{b-a}{n}, \qquad x_j = a + j\,h, \quad j = 0, 1, \ldots, n$$

$$\begin{aligned}
I(f) &= \int_{x_0}^{x_n} f(x)\,dx \\
&= \int_{x_0}^{x_1} f(x)\,dx + \int_{x_1}^{x_2} f(x)\,dx + \cdots + \int_{x_{n-1}}^{x_n} f(x)\,dx \\
T_n(f) &= \frac{h}{2}\left[f(x_0) + f(x_1)\right] + \frac{h}{2}\left[f(x_1) + f(x_2)\right] \\
&\quad + \cdots + \frac{h}{2}\left[f(x_{n-1}) + f(x_n)\right]
\end{aligned}$$

Then the error

$$E_n^T(f) \equiv \int_a^b f(x)\,dx - T_n(f)$$

can be analyzed by adding together the errors over the subintervals $[x_0, x_1]$, $[x_1, x_2]$, ..., $[x_{n-1}, x_n]$. Recall

$$\int_\alpha^{\alpha+h} f(x)\,dx - \frac{h}{2}\left[f(\alpha) + f(\alpha+h)\right] = -\frac{h^3}{12} f''(c)$$

Then on $[x_{j-1}, x_j]$,

$$\int_{x_{j-1}}^{x_j} f(x)\,dx - \frac{h}{2}\left[f(x_{j-1}) + f(x_j)\right] = -\frac{h^3}{12} f''(\gamma_j)$$

with $x_{j-1} \le \gamma_j \le x_j$. Combining these errors, we obtain

$$E_n^T(f) = -\frac{h^3}{12} f''(\gamma_1) - \cdots - \frac{h^3}{12} f''(\gamma_n)$$

This formula can be further simplified, and we will do so in two ways.

Rewrite this error as

$$E_n^T(f) = -\frac{h^3 n}{12}\left[\frac{f''(\gamma_1) + \cdots + f''(\gamma_n)}{n}\right]$$

Denote the quantity inside the brackets by ζ_n. This number satisfies

$$\min_{a \le x \le b} f''(x) \le \zeta_n \le \max_{a \le x \le b} f''(x)$$

Since $f''(x)$ is a continuous function (by original assumption), we have some number c_n in $[a, b]$ for which

$$f''(c_n) = \zeta_n$$

Recall also that $hn = b - a$. Then

$$\begin{aligned} E_n^T(f) &= -\frac{h^3 n}{12}\left[\frac{f''(\gamma_1) + \cdots + f''(\gamma_n)}{n}\right] \\ &= -\frac{h^2 (b-a)}{12} f''(c_n) \end{aligned}$$

This is the error formula given on the first slide.

AN ERROR ESTIMATE

We now obtain a way to estimate the error $E_n^T(f)$. Return to the formula

$$E_n^T(f) = -\frac{h^3}{12} f''(\gamma_1) - \cdots - \frac{h^3}{12} f''(\gamma_n)$$

and rewrite it as

$$E_n^T(f) = -\frac{h^2}{12} \left[f''(\gamma_1)h + \cdots + f''(\gamma_n)h \right]$$

The quantity

$$f''(\gamma_1)h + \cdots + f''(\gamma_n)h$$

is a *Riemann sum* for the integral

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

By this we mean

$$\lim_{n\to\infty} \left[f''(\gamma_1)h + \cdots + f''(\gamma_n)h \right] = \int_a^b f''(x)\,dx$$

Thus

$$f''(\gamma_1)h + \cdots + f''(\gamma_n)h \approx f'(b) - f'(a)$$

for larger values of n. Combining this with the earlier error formula

$$E_n^T(f) = -\frac{h^2}{12}\left[f''(\gamma_1)h + \cdots + f''(\gamma_n)h\right]$$

we have

$$E_n^T(f) \approx -\frac{h^2}{12}\left[f'(b) - f'(a)\right] \equiv \widetilde{E}_n^T(f)$$

This is a computable estimate of the error in the numerical integration. It is called an *asymptotic error estimate*.

Example

Consider evaluating

$$I(f) = \int_0^{\pi} e^x \cos x \, dx = -\frac{e^{\pi}+1}{2} \doteq -12.0703463163896$$

In this case,

$$f'(x) = e^x \left(\cos x - \sin x\right), \quad f''(x) = -2e^x \sin x$$

$$\max_{0 \le x \le \pi} \left| f''(x) \right| = \left| f''\left(.75\pi\right) \right| = 14.921$$

From

$$E_n^T(f) = -\frac{h^2 (b-a)}{12} f''(c_n)$$

we have the *error bound*

$$\left| E_n^T(f) \right| \le \frac{h^2 \pi}{12} \cdot 14.921 = 3.906 h^2 \qquad (2)$$

Also, we have the *error estimate*

$$\widetilde{E}_n^T(f) = -\frac{h^2}{12}\left[f'(\pi) - f'(0)\right] = \frac{h^2}{12}\left(e^{\pi}+1\right) \doteq 2.012 h^2 \qquad (3)$$

n	T_n	E_n^T	Ratio	Bound (2)	$\widetilde{E}_n^T$
2	-17.38925933	5.319		9.638	4.964
4	-13.33602285	1.266	4.20	2.409	1.241
8	-12.38216243	3.118E−1	4.06	6.024E−1	3.103E−1
16	-12.14800410	7.766E−2	4.02	1.506E−1	7.757E−2
32	-12.08974212	1.940E−2	4.00	3.765E−2	1.939E−2
64	-12.07519410	4.848E−3	4.00	9.412E−3	4.848E−3
128	-12.07155818	1.212E−3	4.00	2.353E−3	1.212E−3
256	-12.07064928	3.030E−4	4.00	5.822E−4	3.030E−4

In looking at the table, we see that the error $E_n^T(f)$ and the error estimate $\widetilde{E}_n^T(f)$ are quite close. Therefore

$$I(f) - T_n(f) \approx \frac{h^2}{12}\left(e^\pi + 1\right)$$

$$I(f) \approx T_n(f) + \frac{h^2}{12}\left(e^\pi + 1\right)$$

This last formula is called the *corrected trapezoidal rule*, $CT_n(f)$.

The corrected trapezoidal rule is illustrated in the following table.

n	$I - T_n$	Ratio	$I - CT_n$	Ratio
2	5.319		3.552E−1	
4	1.266	4.20	2.474E−2	14.4
8	3.118E−1	4.06	1.583E−3	15.6
16	7.766E−2	4.02	9.949E−5	15.9
32	1.940E−2	4.00	6.227E−6	16.0
64	4.848E−3	4.00	3.893E−7	16.0
128	1.212E−3	4.00	2.433E−8	16.0
256	3.030E−4	4.00	1.521E−9	16.0

The corrected trapezoidal rule

In general,

$$I(f) - T_n(f) \approx -\frac{h^2}{12}\left[f'(b) - f'(a)\right]$$

$$I(f) \approx CT_n(f) := T_n(f) - \frac{h^2}{12}\left[f'(b) - f'(a)\right]$$

This is the *corrected trapezoidal rule*. It is easy to obtain from the trapezoidal rule, and in most cases, it converges more rapidly than the trapezoidal rule.

ERROR FORMULA FOR SIMPSON'S RULE

Recall the general Simpson's rule

$$\begin{aligned}\int_a^b f(x)\,dx \approx S_n(f) := \frac{h}{3}\,[f(x_0) &+ 4f(x_1) + 2f(x_2) \\ &+ 4f(x_3) + 2f(x_4) + \cdots \\ &+ 2f(x_{n-2}) + 4f(x_{n-1}) + f(x_n)]\end{aligned}$$

For its error, we have

$$E_n^S(f) \equiv \int_a^b f(x)\,dx - S_n(f) = -\frac{h^4\,(b-a)}{180} f^{(4)}(c_n)$$

where $a \le c_n \le b$. For an asymptotic error estimate,

$$\int_a^b f(x)\,dx - S_n(f) \approx \widetilde{E}_n^S(f) \equiv -\frac{h^4}{180}\left[f'''(b) - f'''(a)\right]$$

DISCUSSION

For Simpson's error formula, both formulas assume that the integrand $f(x)$ has four continuous derivatives on the interval $[a, b]$. What happens when this is not valid? We return later to this question.

Both formulas also say the error should decrease by a factor of around 16 when n is doubled.

Compare these results with those for the trapezoidal rule error formulas:

$$E_n^T(f) \equiv \int_a^b f(x)\,dx - T_n(f) = -\frac{h^2(b-a)}{12} f''(c_n)$$

$$E_n^T(f) \approx -\frac{h^2}{12}\left[f'(b) - f'(a)\right] \equiv \widetilde{E}_n^T(f)$$

EXAMPLE

Consider evaluating

$$I = \int_0^2 \frac{dx}{1+x^2}$$

using Simpson's rule $S_n(f)$. Let us bound the error. Begin by noting that

$$f^{(4)}(x) = 24\frac{5x^4 - 10x^2 + 1}{\left(1+x^2\right)^5}$$

$$\max_{0\le x\le 1}\left|f^{(4)}(x)\right| = f^{(4)}(0) = 24$$

Then

$$E_n^S(f) = -\frac{h^4\,(b-a)}{180} f^{(4)}(c_n)$$

$$\left|E_n^S(f)\right| \le \frac{h^4 \cdot 2}{180} \cdot 24 = \frac{4h^4}{15}$$

How large should n be chosen in order to ensure that

$$\left| E_n^S(f) \right| \leq 5 \times 10^{-6}$$

This is true if

$$\begin{aligned} \frac{4h^4}{15} &\leq 5 \times 10^{-6} \\ h &\leq .0658 \\ n &\geq 30.39 \end{aligned}$$

Therefore, choosing $n \geq 32$ will give the desired error bound. Compare this with the earlier trapezoidal example in which $n \geq 517$ was needed.

For the asymptotic error estimate, we have

$$f'''(x) = -24x \frac{x^2 - 1}{(1 + x^2)^4}$$

$$\widetilde{E}_n^S(f) \equiv -\frac{h^4}{180} \left[f'''(2) - f'''(0) \right] = \frac{h^4}{180} \cdot \frac{144}{625} = \frac{4}{3125} h^4$$

INTEGRATING $\sqrt{x}$

Consider the numerical approximation of

$$\int_0^1 \sqrt{x}\,dx = \frac{2}{3}$$

In the following table, we give the errors when using both the trapezoidal and Simpson rules.

n	E_n^T	*Ratio*	E_n^S	*Ratio*
2	$6.311E-2$		$2.860E-2$	
4	$2.338E-2$	2.70	$1.012E-2$	2.82
8	$8.536E-3$	2.74	$3.587E-3$	2.83
16	$3.085E-3$	2.77	$1.268E-3$	2.83
32	$1.108E-3$	2.78	$4.485E-4$	2.83
64	$3.959E-4$	2.80	$1.586E-4$	2.83
128	$1.410E-4$	2.81	$5.606E-5$	2.83

The rate of convergence is slower because the function $f(x) = \sqrt{x}$ is not sufficiently differentiable on $[0,1]$. Both methods converge with a rate proportional to $h^{1.5}$.

ASYMPTOTIC ERROR FORMULAS

For a numerical integration formula,

$$\int_a^b f(x)\,dx \approx \sum_{j=0}^{n} w_j f(x_j)$$

let $E_n(f)$ denote its error,

$$E_n(f) = \int_a^b f(x)\,dx - \sum_{j=0}^{n} w_j f(x_j)$$

We say another formula $\widetilde{E}_n(f)$ is an asymptotic error formula this numerical integration if it satisfies

$$\lim_{n\to\infty} \frac{\widetilde{E}_n(f)}{E_n(f)} = 1 \quad \text{or} \quad \lim_{n\to\infty} \frac{E_n(f) - \widetilde{E}_n(f)}{E_n(f)} = 0$$

These conditions say that $\widetilde{E}_n(f)$ looks increasingly like $E_n(f)$ as n increases, and thus $E_n(f) \approx \widetilde{E}_n(f)$.

For the trapezoidal rule, assuming $f(x)$ has two continuous derivatives on $[a, b]$,

$$E_n^T(f) \approx \widetilde{E}_n^T(f) \equiv -\frac{h^2}{12}\left[f'(b) - f'(a)\right]$$

For Simpson's rule, assuming $f(x)$ has four continuous derivatives on $[a, b]$,

$$E_n^S(f) \approx \widetilde{E}_n^S(f) \equiv -\frac{h^4}{180}\left[f'''(b) - f'''(a)\right]$$

Example (cont.)

Note that both of these formulas can be written in an equivalent form as

$$\widetilde{E}_n(f) = \frac{c}{n^p}$$

for appropriate constant c and exponent p. With the trapezoidal rule, $p = 2$ and

$$c = -\frac{(b-a)^2}{12}\left[f'(b) - f'(a)\right]$$

and for Simpson's rule, $p = 4$ and

$$c = -\frac{(b-a)^4}{180}\left[f'''(b) - f'''(a)\right]$$

The formula

$$\widetilde{E}_n(f) = \frac{c}{n^p} \tag{4}$$

occurs for many other numerical integration formulas that we have not yet defined or studied. In addition, if we use the trapezoidal or Simpson rules with an integrand $f(x)$ which is not sufficiently differentiable, then (4) may hold with a p less than the ideal.

Example. Consider

$$I = \int_0^1 x^\beta f(x)\, dx$$

in which $0 < \beta < 1$ and $f(x)$ is a smooth function with $f(0) \neq 0$. Then the convergence of the trapezoidal rule can be shown to have an asymptotic error formula

$$E_n \approx \widetilde{E}_n = \frac{c}{n^{\beta+1}} \tag{5}$$

for some constant c dependent on β. A similar result holds for Simpson's rule, with $0 < \beta < 3$, β not an integer. We can actually specify a formula for c; but the formula is often less important than knowing that (4) is valid for some c.

Assume we know that an asymptotic error formula

$$I - I_n \approx \frac{c}{n^p}$$

is valid for some numerical integration rule denoted by I_n. Initially, assume we know the exponent p. Then imagine calculating both I_n and I_{2n}. With I_{2n}, we have

$$I - I_{2n} \approx \frac{c}{2^p n^p}$$

This leads to

$$\begin{aligned} I - I_n &\approx 2^p \left(I - I_{2n}\right) \\ I &\approx \frac{2^p I_{2n} - I_n}{2^p - 1} = I_{2n} + \frac{I_{2n} - I_n}{2^p - 1} \end{aligned}$$

The formula

$$I \approx I_{2n} + \frac{I_{2n} - I_n}{2^p - 1} \tag{6}$$

is called *Richardson's extrapolation formula*.

Examples

With the trapezoidal rule and with the integrand $f(x)$ having two continuous derivatives, $p = 2$ and

$$I \approx T_{2n} + \frac{1}{3}\left(T_{2n} - T_n\right)$$

With Simpson's rule and with the integrand $f(x)$ having four continuous derivatives, $p = 4$ and

$$I \approx S_{2n} + \frac{1}{15}\left(S_{2n} - S_n\right)$$

Richardson's error estimate

We can also use the formula (4)

$$\widetilde{E}_n(f) = \frac{c}{n^p}$$

to obtain error estimation formulas:

$$I - I_{2n} \approx \frac{I_{2n} - I_n}{2^p - 1} \tag{7}$$

This is called *Richardson's error estimate*. For example, with the trapezoidal rule,

$$I - T_{2n} \approx \frac{1}{3}\left(T_{2n} - T_n\right)$$

and for Simpson's rule,

$$I - S_{2n} \approx \frac{1}{15}\left(S_{2n} - S_n\right)$$

AITKEN EXTRAPOLATION

In this case, we again assume

$$I - I_n \approx \frac{c}{n^p}$$

But in contrast to previously, we do not know either c or p.
Imagine computing I_n, I_{2n}, and I_{4n}. Then

$$I - I_n \approx \frac{c}{n^p}, \quad I - I_{2n} \approx \frac{c}{2^p n^p}, \quad I - I_{4n} \approx \frac{c}{4^p n^p}$$

We can directly try to estimate I. Dividing

$$\frac{I - I_n}{I - I_{2n}} \approx 2^p \approx \frac{I - I_{2n}}{I - I_{4n}}$$

Solving for I, we obtain

$$\begin{aligned} (I - I_{2n})^2 &\approx (I - I_n)(I - I_{4n}) \\ I(I_n + I_{4n} - 2I_{2n}) &\approx I_n I_{4n} - I_{2n}^2 \\ I &\approx \frac{I_n I_{4n} - I_{2n}^2}{I_n + I_{4n} - 2I_{2n}} \end{aligned}$$

This can be improved computationally, to avoid loss of significance errors.

$$I \approx I_{4n} + \left(\frac{I_n I_{4n} - I_{2n}^2}{I_n + I_{4n} - 2I_{2n}} - I_{4n} \right)$$
$$= I_{4n} - \frac{(I_{4n} - I_{2n})^2}{(I_{4n} - I_{2n}) - (I_{2n} - I_n)}$$

This is called *Aitken's extrapolation formula*.

To estimate p, we use

$$\frac{I_{2n} - I_n}{I_{4n} - I_{2n}} \approx 2^p$$

To see this, write

$$\frac{I_{2n} - I_n}{I_{4n} - I_{2n}} = \frac{(I - I_n) - (I - I_{2n})}{(I - I_{2n}) - (I - I_{4n})}$$

Then substitute from the following and simplify:

$$I - I_n \approx \frac{c}{n^p}, \quad I - I_{2n} \approx \frac{c}{2^p n^p}, \quad I - I_{4n} \approx \frac{c}{4^p n^p}$$

Example

Consider the following table of numerical integrals.

n	I_n	$I_n - I_{n/2}$	*Ratio*
2	.28451779686		
4	.28559254576	$1.075E-3$	
8	.28570248748	$1.099E-4$	9.78
16	.28571317731	$1.069E-5$	10.28
32	.28571418363	$1.006E-6$	10.62
64	.28571427643	$9.280E-8$	10.84

What is its order of convergence? It appears

$$2^p \doteq 10.84, \qquad p \doteq \log_2 10.84 = 3.44$$

We could now combine this with Richardson's error formula:

$$I - I_n \approx \frac{1}{2^p - 1}\left(I_n - I_{n/2}\right)$$

For example,

$$I - I_{64} \approx \frac{1}{9.84}\left[9.280E-8\right] = 9.43E-9$$

PERIODIC FUNCTIONS

A function $f(x)$ is *periodic* if the following condition is satisfied. There is a smallest real number $\tau > 0$ for which

$$f(x+\tau) = f(x), \qquad -\infty < x < \infty \tag{8}$$

The number τ is called the *period* of the function $f(x)$. The constant function $f(x) \equiv 1$ is also considered periodic, but it satisfies this condition with any $\tau > 0$. Basically, a periodic function is one which repeats itself over intervals of length τ.

The condition (8) implies

$$f^{(m)}(x+\tau) = f^{(m)}(x), \qquad -\infty < x < \infty \tag{9}$$

for the m^{th}-derivative of $f(x)$, provided there is such a derivative. Thus the derivatives are also periodic.

Periodic functions occur very frequently in applications of mathematics, reflecting the periodicity of many phenomena in the physical world.

PERIODIC INTEGRANDS

Consider the special class of integrals

$$I(f) = \int_a^b f(x)\,dx$$

in which $f(x)$ is periodic, with $b - a$ an integer multiple of the period τ for $f(x)$. In this case, the performance of the trapezoidal rule and other numerical integration rules is much better than that predicted by earlier error formulas. To hint at this improved performance, recall

$$\int_a^b f(x)\,dx - T_n(f) \approx \widetilde{E}_n(f) \equiv -\frac{h^2}{12}\left[f'(b) - f'(a)\right]$$

With our assumption on the periodicity of $f(x)$, we have

$$f(a) = f(b), \qquad f'(a) = f'(b)$$

Therefore, $\widetilde{E}_n(f) = 0$ and we should expect improved performance in the convergence behaviour of the trapezoidal sums $T_n(f)$.

If in addition to being periodic on $[a, b]$, the integrand $f(x)$ also has m continous derivatives, then it can be shown that

$$I(f) - T_n(f) = \frac{c_m}{n^m} + \text{ smaller terms}$$

By "smaller terms", we mean terms which decrease to zero more rapidly than n^{-m}.

Thus if $f(x)$ is periodic with $b - a$ an integer multiple of the period τ for $f(x)$, and if $f(x)$ is infinitely differentiable, then the error $I - T_n$ decreases to zero more rapidly than n^{-m} for any $m > 0$. For periodic integrands, the trapezoidal rule is an optimal numerical integration method.

Example

Consider evaluating

$$I = \int_0^{2\pi} \frac{\sin x}{1 + e^{\sin x}}\, dx$$

Using the trapezoidal rule, we have the results in the following table. In this case, the formulas based on Richardson extrapolation are no longer valid.

n	T_n	$T_n - T_{\frac{1}{2}n}$
2	0.0	
4	-0.72589193317292	$-7.259E-1$
8	-0.74006131211583	$-1.417E-2$
16	-0.74006942337672	$-8.111E-6$
32	-0.74006942337946	$-2.746E-12$
64	-0.74006942337946	0.0