NUMERICAL INTEGRATION

How do you evaluate

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

From calculus, if $F(x)$ is an antiderivative of $f(x)$, then

$$I = \int_a^b f(x)\,dx = F(x)|_a^b = F(b) - F(a)$$

However, in practice most integrals cannot be evaluated by this means. And even when this can work, an approximate numerical method may be much simpler and easier to use. For example, the integrand in

$$\int_0^1 \frac{dx}{1 + x^5}$$

has an extremely complicated antiderivative; and it is easier to evaluate the integral by approximate means. Try evaluating this integral with *Maple* or *Mathematica*.

NUMERICAL INTEGRATION
A GENERAL FRAMEWORK

Returning to a lesson used earlier with rootfinding:
If you cannot solve a problem, then replace it with a "near-by" problem that you can solve.
In our case, we want to evaluate

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

To do so, many of the numerical schemes are based on choosing approximates of $f(x)$. Calling one such $\widetilde{f}(x)$, use

$$I \approx \int_a^b \widetilde{f}(x)\,dx \equiv \widetilde{I}$$

What is the error?

$$\begin{aligned}
E &= I - \widetilde{I} = \int_a^b \left[f(x) - \widetilde{f}(x) \right] \, dx \\
|E| &\leq \int_a^b \left| f(x) - \widetilde{f}(x) \right| \, dx \\
&\leq (b-a) \left\| f - \widetilde{f} \right\|_\infty \\
\left\| f - \widetilde{f} \right\|_\infty &\equiv \max_{a \leq x \leq b} \left| f(x) - \widetilde{f}(x) \right|
\end{aligned}$$

We also want to choose the approximates $\widetilde{f}(x)$ of a form we can integrate directly and easily. Examples are polynomials, trig functions, piecewise polynomials, and others.

If we use polynomial approximations, then how do we choose them. At this point, we have two choices:

1. *Taylor polynomials* approximating $f(x)$

2. *Interpolatory polynomials* approximating $f(x)$

EXAMPLE

Consider evaluating

$$I = \int_0^1 e^{x^2}\, dx$$

Use

$$\begin{aligned} e^t &= 1 + t + \tfrac{1}{2!}t^2 + \cdots + \tfrac{1}{n!}t^n + \tfrac{1}{(n+1)!}t^{n+1}e^{c_t} \\ e^{x^2} &= 1 + x^2 + \tfrac{1}{2!}x^4 + \cdots + \tfrac{1}{n!}x^{2n} + \tfrac{1}{(n+1)!}x^{2n+2}e^{d_x} \end{aligned}$$

with $0 \le d_x \le x^2$. Then

$$\begin{aligned} I = \int_0^1 &\left[1 + x^2 + \tfrac{1}{2!}x^4 + \cdots + \tfrac{1}{n!}x^{2n}\right] dx \\ &+ \tfrac{1}{(n+1)!}\int_0^1 \left[x^{2n+2}e^{d_x}\right] dx \end{aligned}$$

Taking $n = 3$, we have

$$\begin{aligned} I &= 1 + \tfrac{1}{3} + \tfrac{1}{10} + \tfrac{1}{42} + E = 1.4571 + E \\ 0 < E &\le \tfrac{e}{24}\int_0^1 \left[x^8\right] dx = \tfrac{e}{216} = .0126 \end{aligned}$$

USING INTERPOLATORY POLYNOMIALS

In spite of the simplicity of the above example, it is generally more difficult to do numerical integration by constructing Taylor polynomial approximations than by constructing polynomial interpolates. We therefore construct the function $\widetilde{f}$ in

$$\int_a^b f(x)\,dx \approx \int_a^b \widetilde{f}(x)\,dx$$

by means of interpolation.

Initially, we consider only the case in which the interpolation is based on interpolation at evenly spaced node points.

LINEAR INTERPOLATION

The linear interpolant to $f(x)$, interpolating at a and b, is given by

$$P_1(x) = \frac{(b-x)\,f(a) + (x-a)\,f(b)}{b-a}$$

Using the linear interpolant

$$P_1(x) = \frac{(b-x)\,f(a) + (x-a)\,f(b)}{b-a}$$

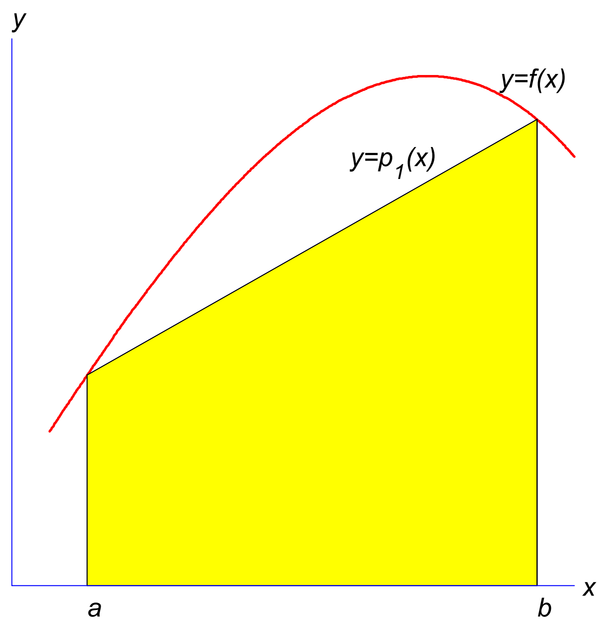
we obtain the approximation

$$\begin{aligned} \int_a^b f(x)\,dx &\approx \int_a^b P_1(x)\,dx \\ &= \tfrac{1}{2}\,(b-a)\,[f(a)+f(b)] \equiv T_1(f) \end{aligned}$$

The rule

$$\int\limits_a^b f(x)\,dx \approx T_1(f)$$

is called the *trapezoidal rule*.



Illustrating $I \approx T_1(f)$

Example.

$$\begin{aligned} \int_0^{\pi/2} \sin x \, dx &\approx \frac{\pi}{4}\left[\sin 0 + \sin\left(\frac{\pi}{2}\right)\right] \\ &= \frac{\pi}{4} \doteq .785398 \\ Error &= .215 \end{aligned}$$

HOW TO OBTAIN GREATER ACCURACY?

How do we improve our estimate of the integral

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

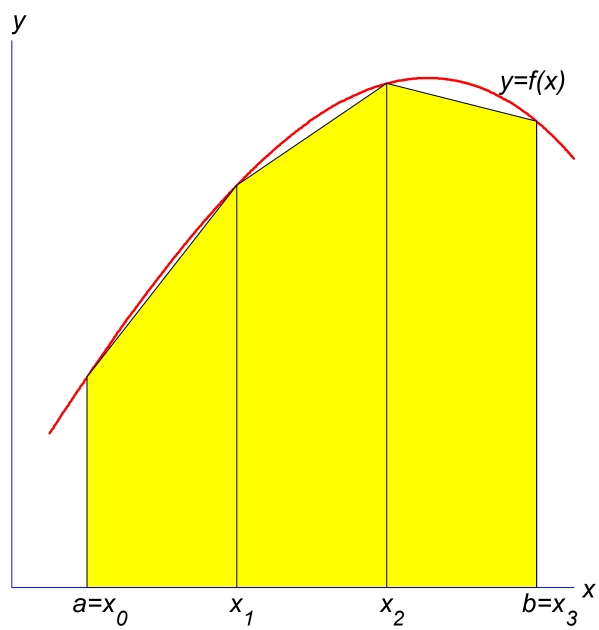
One direction is to increase the degree of the approximation, moving next to a quadratic interpolating polynomial for $f(x)$. We first look at an alternative.

Instead of using the trapezoidal rule on the original interval $[a, b]$, apply it to integrals of $f(x)$ over smaller subintervals. For example:

$$\begin{aligned} I &= \int_a^c f(x)\, dx + \int_c^b f(x)\, dx, \qquad c = \tfrac{b+a}{2} \\ &\approx \tfrac{c-a}{2}\, [f(a) + f(c)] + \tfrac{b-c}{2}\, [f(c) + f(b)] \\ &= \tfrac{h}{2}\, [f(a) + 2f(c) + f(b)] \equiv T_2(f), \qquad h = \tfrac{b-a}{2} \end{aligned}$$

Example.

$$\begin{aligned} \int_0^{\pi/2} \sin x\, dx &\approx \tfrac{\pi}{8} \left[\sin 0 + 2 \sin\left(\tfrac{\pi}{4}\right) + \sin\left(\tfrac{\pi}{2}\right)\right] \\ &\doteq .948059 \\ Error &= .0519 \end{aligned}$$



Illustrating $I \approx T_3(f)$

THE TRAPEZOIDAL RULE

We can continue as above by dividing $[a, b]$ into even smaller subintervals and applying

$$\int_{\alpha}^{\beta} f(x)\,dx \approx \frac{\beta-\alpha}{2}\left[f(\alpha)+f(\beta)\right], \qquad (*)$$

on each of the smaller subintervals. Begin by introducing a positive integer $n \geq 1$,

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

Then

$$\begin{aligned} I &= \int_{x_0}^{x_n} f(x)\,dx \\ &= \int_{x_0}^{x_1} f(x)\,dx + \int_{x_1}^{x_2} f(x)\,dx + \cdots + \int_{x_{n-1}}^{x_n} f(x)\,dx \end{aligned}$$

Use $[\alpha, \beta] = [x_0, x_1], [x_1, x_2], ..., [x_{n-1}, x_n]$, for each of which the subinterval has length h.

Then applying

$$\int_{\alpha}^{\beta} f(x)\,dx \approx \frac{\beta-\alpha}{2}\left[f(\alpha)+f(\beta)\right]$$

we have

$$\begin{aligned} I \approx\ & \tfrac{h}{2}\left[f(x_0)+f(x_1)\right]+\tfrac{h}{2}\left[f(x_1)+f(x_2)\right] \\ & +\cdots \\ & +\tfrac{h}{2}\left[f(x_{n-2})+f(x_{n-1})\right]+\tfrac{h}{2}\left[f(x_{n-1})+f(x_n)\right] \end{aligned}$$

Simplifying,

$$\begin{aligned} I &\approx h\left[\frac{1}{2}f(a)+f(x_1)+\cdots+f(x_{n-1})+\frac{1}{2}f(b)\right] \\ &\equiv T_n(f) \end{aligned}$$

This is called the "composite trapezoidal rule", or more simply, the *trapezoidal rule*.

Example. Again integrate $\sin x$ over $\left[0, \frac{\pi}{2}\right]$. Then we have

n	$T_n(f)$	$Error$	$Ratio$
1	.785398163	2.15E−1	
2	.948059449	5.19E−2	4.13
4	.987115801	1.29E−2	4.03
8	.996785172	3.21E−3	4.01
16	.999196680	8.03E−4	4.00
32	.999799194	2.01E−4	4.00
64	.999949800	5.02E−5	4.00
128	.999987450	1.26E−5	4.00
256	.999996863	3.14E−6	4.00

Note that the errors are decreasing by a constant factor of 4. Why do we always double n?

USING QUADRATIC INTERPOLATION

We want to approximate $I = \int_a^b f(x)\,dx$ using quadratic interpolation of $f(x)$. Interpolate $f(x)$ at the points $\{a, c, b\}$, with $c = \frac{1}{2}(a+b)$. Also let $h = \frac{1}{2}(b-a)$. The quadratic interpolating polynomial is given by

$$\begin{aligned} P_2(x) &= \frac{(x-c)(x-b)}{2h^2} f(a) + \frac{(x-a)(x-b)}{-h^2} f(c) \\ &\quad + \frac{(x-a)(x-c)}{2h^2} f(b) \end{aligned}$$

Replacing $f(x)$ by $P_2(x)$, we obtain the approximation

$$\begin{aligned} \int_a^b f(x)\,dx &\approx \int_a^b P_2(x)\,dx \\ &= \frac{h}{3}\left[f(a) + 4f(c) + f(b)\right] \equiv S_2(f) \end{aligned}$$

This is called *Simpson's rule*.

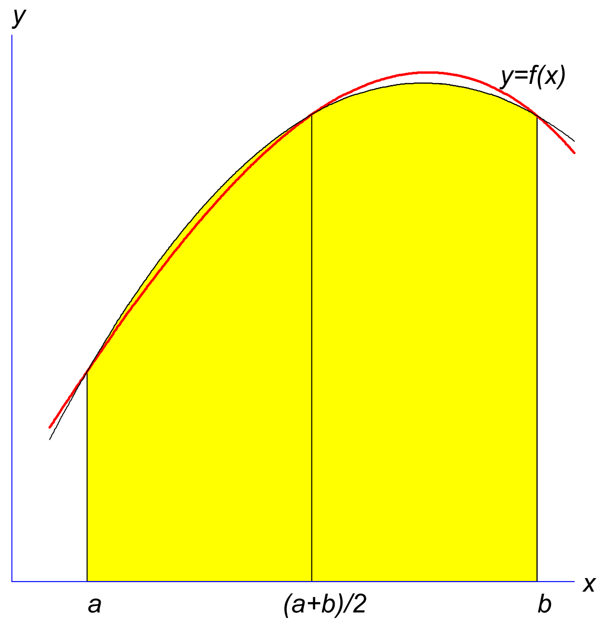


Illustration of $I \approx S_2(f)$

Example.

$$\begin{aligned}
\int_0^{\pi/2} \sin x \, dx &\approx \frac{\pi/2}{3}\left[\sin 0 + 4\sin\left(\frac{\pi}{4}\right) + \sin\left(\frac{\pi}{2}\right)\right] \\
&\doteq 1.00227987749221 \\
Error &= -0.00228
\end{aligned}$$

SIMPSON'S RULE

As with the trapezoidal rule, we can apply Simpson's rule on smaller subdivisions in order to obtain better accuracy in approximating

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

Again, Simpson's rule is given by

$$\int_\alpha^\beta f(x)\,dx \approx \frac{\delta}{3}\left[f(\alpha) + 4f(\gamma) + f(\beta)\right], \quad \gamma = \frac{\alpha+\beta}{2}$$

and $\delta = \frac{1}{2}(\beta - \alpha)$.

Let n be a positive even integer, and

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

Then write

$$\begin{aligned} I &= \int_{x_0}^{x_n} f(x)\,dx \\ &= \int_{x_0}^{x_2} f(x)\,dx + \int_{x_2}^{x_4} f(x)\,dx + \cdots + \int_{x_{n-2}}^{x_n} f(x)\,dx \end{aligned}$$

Apply

$$\int_{\alpha}^{\beta} f(x)\,dx \approx \frac{\delta}{3}\left[f(\alpha) + 4f(\gamma) + f(\beta)\right], \quad \gamma = \frac{\alpha+\beta}{2}$$

to each of these subintegrals, with

$$[\alpha, \beta] = [x_0, x_2], [x_2, x_4], ..., [x_{n-2}, x_n]$$

In all cases, $\frac{1}{2}(\beta - \alpha) = h$. Then

$$\begin{aligned} I \approx\ & \tfrac{h}{3}\left[f(x_0) + 4f(x_1) + f(x_2)\right] \\ & + \tfrac{h}{3}\left[f(x_2) + 4f(x_3) + f(x_4)\right] \\ & + \cdots \\ & + \tfrac{h}{3}\left[f(x_{n-4}) + 4f(x_{n-3}) + f(x_{n-2})\right] \\ & + \tfrac{h}{3}\left[f(x_{n-2}) + 4f(x_{n-1}) + f(x_n)\right] \end{aligned}$$

This can be simplified to

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

This is called the "composite Simpson's rule" or more simply, .*Simpson's rule*

EXAMPLE

Approximate $\int_0^{\pi/2} \sin x \, dx$. The Simpson rule results are as follows.

n	$S_n(f)$	$Error$	$Ratio$
2	1.00227987749221	−2.28E−3	
4	1.00013458497419	−1.35E−4	16.94
8	1.00000829552397	−8.30E−6	16.22
16	1.00000051668471	−5.17E−7	16.06
32	1.00000003226500	−3.23E−8	16.01
64	1.00000000201613	−2.02E−9	16.00
128	1.00000000012600	−1.26E−10	16.00
256	1.00000000000788	−7.88E−12	16.00
512	1.00000000000049	−4.92E−13	15.99

Note that the ratios of successive errors have converged to 16. Why? Also compare this table with that for the trapezoidal rule. For example,

$$\begin{aligned} I - T_4 &= 1.29E-2 \\ I - S_4 &= -1.35E-4 \end{aligned}$$

There is a great deal to be learned by doing numbers of examples. For example, are the ratios of convergence for our numerical examples typical of the trapezoidal and Simpson rules? Several of these are given on pages 188 (for trapezoidal rule) and 192 (for Simpson's rule). They are for the integrals

$$
\begin{aligned}
I^{(1)} &= \int_0^1 e^{-x^2}dx \doteq .74682413281234 \\
I^{(2)} &= \int_0^4 \frac{dx}{1+x^2} = \arctan 4 \\
I^{(3)} &= \int_0^{2\pi} \frac{dx}{2+\cos x} = \frac{2\pi}{\text{sqrt}(3)}
\end{aligned}
$$

Look carefully at those tables.