Romberg Integration

We have already seen the Composite Trapezoid Rule: If $f \in C^2[a,b]$ then there exists a $\mu \in (a,b)$ so that

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

where $h = \frac{b-a}{n}$, $x_j = a + jh$

Thus the error for the Composite Trapezoid Rule is $O(h^2)$

In fact we can be more precise. An application of the Euler-MacLauren Summation Formula shows that for sufficiently smooth f

$$\text{error} = C_1 h^2 + C_2 h^4 + \cdots + C_m h^{2m} + O(h^{2m+2})$$

where $C_K = \text{const} \times \left(f^{(2K-1)}(b) - f^{(2K-1)}(a)\right)$

This shows us that the Composite Trapezoid Rule is extremely accurate for smooth periodic functions provided h is small enough.

Ex The numerical approximation to

$$\int_0^1 \sin^2(8\pi x)\,dx$$

by the composite trapezoid rule for several values of h is given below

h	1	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	$\cdots$
Composite Trapezoid Rule	0	0	0	0	$\frac{1}{2}$	$\frac{1}{2}$	$\cdots$

Notice also that we know the form of the error, so we can obtain higher order accuracy by using Richardson extrapolation. (To give Romberg Integration)

We will carry out Composite Trapezoid Rule approximations with

$$m_1 = 1,\ m_2 = 2,\ m_3 = 4,\ \ldots,\ \text{and } m_n = 2^{n-1}$$

intervals.

The values of the step sizes h_k corresponding to m_k are

$$h_k = (b-a)/m_k = (b-a)/2^{k-1}$$

With this notation, the Composite Trapezoid Rule becomes

$$\int_a^b f(x)\,dx = \frac{h_k}{2}\left[f(a) + f(b) + 2\left(\sum_{i=1}^{2^{k-1}} f(a + ih_k)\right)\right) - \frac{(b-a)}{12} h_k^2 f''(\mu_k)$$

where $\mu_k \in (a, b)$.

Let $R_{k,1}$ be the approximation to the integral using $m_k = 2^{k-1}$ intervals.

ie $\quad R_{1,1} = \frac{h_1}{2}\left[f(a) + f(b)\right] = \frac{(b-a)}{2}\left[f(a) + f(b)\right]$

$$R_{2,1} = \frac{h_2}{2}\left[f(a) + f(b) + 2f(a + h_2)\right]$$

$$= \frac{(b-a)}{4}\left[f(a) + f(b) + 2f\left(a + \frac{b-a}{2}\right)\right]$$

$$= \frac{1}{2}\left[R_{1,1} + h_1 f(a + h_2)\right]$$

notice that when h is halved all the old points at which the function was evaluated appear in the new computation & we thus can avoid repeating the evaluations.

$$R_{3,1} = \frac{1}{2}\left\{R_{2,1} + h_2\left[f(a+h_3) + f(a+3h_3)\right]\right\}$$

$$\vdots$$

$$R_{K,1} = \frac{1}{2}\left\{R_{K-1,1} + h_{K-1}\sum_{i=1}^{2^{K-2}} f(a+(2i-1)h_K)\right\}$$

We can apply this equation to perform the first step of Romberg integration for

$$\int_0^1 e^{-x}\,dx = 1 - e^{-1} \doteq 0.63212$$

$$R_{1,1} = \frac{(1-0)}{2}\left[e^{-0} + e^{-1}\right] \approx 0.68394$$

$$R_{2,1} = \frac{1}{2}\left[R_{1,1} + \frac{1-0}{2} e^{-(0+\frac{1}{2})}\right] \simeq 0.64523$$

$$R_{3,1} = 0.63541$$

$$R_{4,1} = 0.63294$$

We can obtain a faster convergence using Richardson Extrapolation:

Notice that

$$\int_a^b f(x)\,dx - R_{K,1} = \sum_{i=1}^{m} c_i h_K^{2i} + O(h^{2m+2})$$
$$= c_1 h_K^2 + \sum_{i=2}^{m} c_i h_K^{2i} + O(h^{2m+2})$$

$$\int_a^b f(x)\,dx - R_{K+1,1} = \sum_{i=1}^{m} c_i h_{K+1}^{2i} + O(h^{2m+2})$$
$$= \frac{c_1}{4} h_K^2 + \sum_{i=2}^{m} \left(\frac{c_i h_K^{2i}}{4^i}\right) + O(h^{2m+2})$$

Subtracting the first from 4 times the second gives an $O(h_K^4)$ formula:

$$\int_a^b f(x)\,dx - R_{K,2} = \sum_{i=2}^{m} \frac{c_i}{3}\left(\frac{h_K^{2i}}{4^{i-1}} - h_K^{2i}\right) + O(h^{2m+2})$$

where $R_{K,2} = R_{K,1} + \dfrac{R_{K,2} - R_{K-1,1}}{3}$

Of course, this procedure can be repeated to eliminate the $O(h_K^4)$ term from the error.

Continuing in this manner, we have an $O(h_K^{2j})$ approximation formula defined by

$$R_{K,j} = R_{K,j-1} + \frac{R_{K,j-1} - R_{K-1,j-1}}{4^{j-1} - 1}$$

Ex Use Romberg Integration to approximate

$$\int_0^1 e^{-x}\,dx$$

to 5 significant digits

Ans

	$R_{K,1}$	$R_{K,2}$	$R_{K,3}$	$R_{K,4}$	$R_{K,5}$
$R_{1,j}$	.6839397				
$R_{2,j}$	.6452352	.6323337			
$R_{3,j}$	.6354094	.6321342	.6321209		
$R_{4,j}$	.6329434	.6321214	.6321206	.6321206	
$R_{5,j}$	.6323263	.6321206	.6321206	.6321206	.6321206

A typical stopping criteria is that both

$$|R_{n-1,n-1} - R_{n,n}|$$

and $$|R_{n-2,n-2} - R_{n-1,n-1}|$$

are within a desired tolerance.

Note that we may not observe the expected convergence acceleration if

- The integrand f is not sufficiently smooth. We need $f \in C^{2k+2}[a,b]$ to generate the k^{th} row of the table.
- The coefficients $c_1, c_2, \ldots$ are very small. This happens for periodic functions if the interval of integration is an integer multiple of the period or for functions with extremely small derivatives at the endpoints of the interval of integration.

Adaptive Quadrature.

Composite quadrature rules necessitate the use of equally spaced points.

This does not take into account that some portions of the curve may have lange functional variations that require more attention than other portions of the curve.

It is useful to introduce a method that adjusts the step size to be smaller over portions of the curve where a larger functional variation occurs. This technique is called adaptive quadrature.

We will now discuss an adaptive quadrature based on Simpson's Rule. The other composite procedures can be modified in a similar manner.

We want to approximate

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

to within a specified tolerance $\varepsilon > 0$.

Start by applying Simpson's Rule with a step size $h = (b-a)/2$

$$\int_a^b f(x)\,dx = S(a,b) - \frac{h^5}{90} f^{(4)}(\mu) \qquad (*)$$

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

μ is some constant, $a < \mu < b$

We want to know if we should further subdivide the interval, so we need an estimate for the error.

Unfortunately, we don't know $f^{(4)}(\mu)$.

Instead, we will estimate the error using Simpson's Rule with a step size $(b-a)/4$:

$$\int_a^b f(x)\,dx = \frac{h}{6}\Big[f(a) + 4f\left(a+\tfrac{h}{2}\right) + 2f(a+h) + 4f\left(a+\tfrac{3h}{2}\right) + f(b)\Big] - \left(\frac{h}{2}\right)^4 \frac{(b-a)}{180} f^{(4)}(\tilde{\mu})$$

for some $\tilde{\mu} \in (a,b)$.

Or

$$\int_a^b f(x)\,dx = S\left(a, \tfrac{a+b}{2}\right) + S\left(\tfrac{a+b}{2}, b\right) - \frac{1}{16}\frac{h^5}{90} f^{(4)}(\tilde{\mu}), \quad (**)$$

where $S\left(a, \frac{a+b}{2}\right) = \frac{h}{6}\left[f(a) + 4f\left(a+\frac{h}{2}\right) + f(a+h)\right]$

and $S\left(\frac{a+b}{2}, b\right) = \frac{h}{6}\left[f(a+h) + 4f\left(a+\frac{3h}{2}\right) + f(b)\right]$

Now we must make an assumption:

$$\boxed{\text{We assume that } f^{(4)}(\mu) \approx f^{(4)}(\tilde{\mu})}$$

<u>Note</u> if $f^{(4)}$ is continuous, this assumption will hold for sufficiently small h.

Subtract (**) from (*):

$$0 \doteq S(a,b) - S\left(a, \tfrac{a+b}{2}\right) - S\left(\tfrac{a+b}{2}, b\right) - \tfrac{15}{16}\tfrac{h^5}{90} f^{(4)}(\mu)$$

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

$$\therefore \ |\text{error}| = \left| \int_a^b f(x)\,dx - S\left(a, \tfrac{a+b}{2}\right) - S\left(\tfrac{a+b}{2}, b\right)\right|$$
$$\doteq \frac{1}{16}\left|\frac{h^5}{90} f^{(4)}(\mu)\right|$$
$$\doteq \frac{1}{15}\left|\underbrace{S(a,b) - S\left(a, \tfrac{a+b}{2}\right) - S\left(\tfrac{a+b}{2}, b\right)}_{\text{these are all known quantities.}}\right|$$

Thus if we want $|\text{error}| < \varepsilon$ we typically insist that

$$\frac{1}{10}\left|S(a,b) - S\left(a, \tfrac{a+b}{2}\right) - S\left(\tfrac{a+b}{2}, b\right)\right| < \varepsilon$$

Often a factor $\frac{1}{10}$ is used rather that $\frac{1}{15}$ since we had to make the assumption $f^{(4)}(\mu) \approx f^{(4)}(\tilde{\mu})$ and we prefer to be somewhat conservative in our error estimate.

From last day, we had an error estimate for Simpson's Rule:

$$\text{error estimate} = \frac{1}{15}\left|S(a,b) - S(a, \tfrac{a+b}{2}) - S(\tfrac{a+b}{2}, b)\right|$$

<u>Ex</u> Compute the Simpson's Rule approximations $S(a,b)$, $S(a, \frac{a+b}{2})$, $S(\frac{a+b}{2}, b)$ for

$$\int_1^{1.5} \underbrace{x^2 \ln(x)}_{f(x)}\, dx = 0.19225935$$

and calculate the error estimate and the actual error.

<u>Ans</u>: $S(1, 1.5) = \frac{.25}{3}\left[f(1) + 4f(1.25) + f(1.5)\right]$

$= 0.19224530$

$S(1, 1.25) = 0.03937 2434$

$S(1.25, 1.5) = 0.15288602$

$\therefore$ error estimate $= \frac{1}{15}\left|S(1,1.5) - S(1,1.25) - S(1.25,1.5)\right|$

$\approx 8.77 \times 10^{-7}$

actual error $= \left|S(1,1.25) + S(1.25,1.5) - 0.19225935\right|$

$\approx 8.96 \times 10^{-7}$

In deriving our error estimate, we had to make the assumption that

$$f^{(4)}(\mu) \approx f^{(4)}(\tilde{\mu})$$

and to compensate for this error we typically insist that

$$\frac{1}{15}\left| S(a,b) - S\left(a, \tfrac{a+b}{2}\right) - S\left(\tfrac{a+b}{2}, b\right) \right| < \underbrace{\tfrac{2}{3}\varepsilon}_{\equiv \tilde{\varepsilon}}.$$

Adaptive quadradure methods are built on this condition.

If the error estimate is less than $\frac{2}{3}$ the desired tolerance ε then

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

is assumed to be a sufficiently accurate approximation to $\int_a^b f(x)\,dx$.

Otherwise we apply Simpson's Rule to the subintervals $\left[a, \frac{a+b}{2}\right]$ and $\left[(a+b)/2, b\right]$. We also use the error estimation in each of the subintervals. If each error estimate is less than $\varepsilon/2$, then we sum the approximations to get an approximation of $\int_a^b f(x)\,dx$.

If the approximation on one of the subintervals fails to be within the tolerance $\tilde{\varepsilon}/2$ then that subinterval is itself subdivided and the procedure is reapplied to the two subintervals to determine if the approximation on each subinterval is accurate to within $\tilde{\varepsilon}/4$.

This recursive procedure is continued until each portion is within the desired tolerance.

Gaussian Quadrature

Thus far we have only dealt with quadrature formulae

$$\int_a^b f(x)\,dx \approx \sum_{j=1}^{n} a_j\, f(x_j)$$

that relied on nodes that are equally spaced. This is a nice feature for composite rules because it reduces the number of function evaluations.

If we allow ourselves to use unequally spaced points, however, then we can construct quadrature formulas of higher order accuracy.

Notice that with n nodes and n weights we have $2n$ free parameters & we may hope to find an optimal quadrature formula which is exact for polynomials of degree $\leq 2n-1$.

First, note that an integral

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

over an interval $[a, b]$ can be transformed into an integral over $[-1, 1]$ using a change of variables.

Let $$t = \frac{2x - a - b}{b - a}$$

then $$x = \tfrac{1}{2}\left[(b-a)t + a + b\right]$$

and $$dx = \tfrac{1}{2}(b-a)\,dt.$$

So

$$\int_a^b f(x)\,dx = \int_{-1}^{1} f\left(\tfrac{1}{2}\left[(b-a)t + a + b\right]\right) \tfrac{1}{2}(b-a)\,dt$$

So, without loss of generality, we will consider integrals over the interval $[-1, 1]$.

Suppose that $n=2$ (2 nodes) and that we want to determine c_1, c_2, x_1, x_2 so that the integration formula

$$\int_{-1}^{1} f(x)\,dx = c_1 f(x_1) + c_2 f(x_2)$$

gives the exact result whenever $f(x)$ is a polynomial of degree $2(2)-1=3$!

i.e. $f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3$

Since $\int_{-1}^{1} a_0 + a_1 x + a_2 x^2 + a_3 x^3\,dx$

$$= a_0\int_{-1}^{1} dx + a_1\int_{-1}^{1} x\,dx + a_2\int_{-1}^{1} x^2\,dx + a_3\int_{-1}^{1} x^3\,dx$$

this problem is equivalent to showing that the formula is exact for $f(x) = 1, x, x^2, x^3$.

CASES:

$f(x)=1$: $c_1 + c_2 = a_0\int_{-1}^{1} dx$

$f(x)=x$: $c_1x_1 + c_2x_2 = \int_{-1}^{1} x\,dx$

$f(x)=x^2$: $c_1x_1^2 + c_2x_2^2 = \int_{-1}^{1} x^2\,dx$

$f(x)=x^3$: $c_1x_1^3 + c_2x_2^3 = \int_{-1}^{1} x^3\,dx$

Solving this system gives us that

$$c_1 = 1,\ c_2 = 1,\ x_1 = -\sqrt{3}/3,\ x_2 = \sqrt{3}/3$$

$\Rightarrow$ the approximation formula is

$$\int_{-1}^{1} f(x)\,dx = f(-\sqrt{3}/3) + f(\sqrt{3}/3).$$

This approach can be used to obtain the nodes and coefficients for larger n, but Legendre polynomials can be used to obtain them more easily.

The Legendre polynomials $P_0(x), P_1(x), \ldots$ are defined according to the following 2 properties

1. $P_n(x)$ is a polynomial of degree n
2. $\int_{-1}^{1} P(x) P_n(x)\,dx = 0$ whenever $P(x)$ is a polynomial of degree less than n.

The first few Legendre polynomials are

$P_0(x) = 1$, $P_1(x) = x$, $P_2(x) = x^2 - \frac{1}{3}$, $P_3(x) = x^3 - \frac{3}{5}x$, $P_4(x) = x^4 - \frac{6}{7}x^2 + \frac{3}{35}$

Some properties:

- the roots of these polynomials are distinct
- The roots of these polynomials lie in $(-1, 1)$
- The P_n's are symmetrical about the origin $\Rightarrow$ the roots are symmetrical about the origin
- The roots of the n^{th} degree Legendre polynomial have the property that they are the nodes needed to produce an integral approximation formula that gives the exact results for any polynomial of degree less than $2n$.

Thm Suppose $x_1, x_2, \dots, x_n$ are the roots of the n^{th} Legendre polynomial $P_n(x)$ and that for each $i = 1, 2, \dots, n$ the numbers c_i are defined by

$$c_i = \int_{-1}^{1} \prod_{\substack{j=1 \\ j \neq i}}^{n} \frac{x - x_j}{x_i - x_j}\, dx$$

If $P(x)$ is any polynomial of degree less than $2n$ then

$$\int_{-1}^{1} P(x)\, dx = \sum_{i=1}^{n} c_i P(x_i)$$

Proof STEP 1: SHOW THAT THE STATEMENT IS TRUE FOR POLYNOMIALS OF DEGREE $< n$.

Suppose $R(x)$ is a polynomial of degree $< n$.

$$\int_{-1}^{1} R(x)\, dx = \int_{-1}^{1} \underbrace{\left[\sum_{i=1}^{n} \prod_{\substack{j=1 \\ j \neq i}}^{n} \frac{x - x_j}{x_i - x_j} R(x_i) \right]}_{} dx$$

just the $(n-1)$st Lagrange interpolating polynomial with nodes at the roots of the n^{th} Legendre polynomial.

$$= \sum_{i=1}^{n} \left[\int_{-1}^{1} \prod_{\substack{j=1 \\ j \neq i}}^{n} \frac{x - x_j}{x_i - x_j} \right] R(x_i) = \sum_{i=1}^{n} c_i R(x_i)$$

which verifies the result for polynomials of degree less than n.

<u>STEP 2</u>: show true for polynomials of degree less than $2n$.

Suppose $P(x)$ is a polynomial of degree $<2n$.

If we divide P by the n^{th} Legendre polynomial $P_n(x)$ we get

$$P(x) = Q(x)\,P_n(x) + R(x)$$

Q & R are of degree $<n$.

Now
$$\int_{-1}^{1} P(x)\,dx = \int_{-1}^{1} Q(x)P_n(x)\,dx + \int_{-1}^{1} R(x)\,dx$$

$\deg(Q)<n$ so this term is 0, by the defn of Legendre polynomials.

By step 1, $\int_{-1}^{1} R(x)\,dx = \sum_{i=1}^{n} c_i R(x_i)$ since $\deg(R)<n$

$$\therefore \int_{-1}^{1} P(x)\,dx = \sum_{i=1}^{n} c_i R(x_i)$$
$$= \sum_{i=1}^{n} c_i \left[P(x_i) - Q(x_i)P_n(x_i)\right]$$
$$= \sum_{i=1}^{n} c_i P(x_i)$$

∎

The c_i's and the nodes are both extensively tabulated (see text for a reference).

Also, high order Legendre polynomials are built into Maple.

```
with(orthopoly);
P(n, x);
```

$\leftarrow$ gives the n^{th} Legendre polynomial

<u>Ex</u> Approximate $\int_1^{3/2} x^2 \ln x \, dx$ using Gaussian quadrature with $n=2$

$$\int_1^{3/2} x^2 \ln x \, dx = \int_{-1}^{1} \left(\frac{\frac{1}{2}t + \frac{5}{2}}{2}\right)^2 \ln\left(\frac{\frac{1}{2}t + 5/2}{2}\right)\left(\frac{1/2}{2}\right) dt$$

$$\therefore \int_1^{3/2} x^2 \ln x \, dx \approx \frac{1}{4}\left(\frac{\frac{1}{2}x_1 + 5/2}{2}\right)^2 \ln\left(\frac{\frac{1}{2}x_1 + 5/2}{2}\right) + \frac{1}{4}\left(\frac{\frac{1}{2}x_2 + 5/2}{2}\right)^2 \ln\left(\frac{\frac{1}{2}x_2 + 5/2}{2}\right)$$

where $x_1 \approx 0.5773502692$
$x_2 \approx -x_1$

Thus $\int_1^{3/2} x^2 \ln x \, dx \approx 0.1922687$

(Recall that Simpson's Rule gave 0.19224530 for this problem & that the exact result is 0.19225935).