Simpson's Rule, fine details

Richardson Extrapolation

Recall that the error in the composite Simpson's Rule is

$$\varepsilon = (b-a)h^4 f^{(4)}(\xi)/180,$$

where h is the strip width and ξ is some number between a and b. What happens when we go from I_2 to I_4 to I_8 [and, if necessary, I_{16}, and so on]? Well, h is halved, so h^4 is divided by 16; but everything else stays much the same. [It *looks* exactly the same, but ξ will usually be a different number.] So the error is *approximately* divided by 16. We can use this to estimate how big the error is! For example, in our standard integral, $I = \int_1^2 1/x \, \mathrm{d}x$, we had $I_4 \approx 0.693254$ and $I_8 \approx 0.693155$. So if ε is the error in the last result, we have:

$$I = I_8 - \varepsilon; \; I \approx I_4 - 16 \times \varepsilon.$$

So, subtracting,

$$I_8 - I_4 \approx 0.693155 - 0.693254 = -0.000099 \approx -15\varepsilon,$$

and so $\varepsilon \approx 0.000007$ and $I \approx 0.693148$, compared with $\log 2 \approx 0.693147$.

In return for some essentially trivial arithmetic, $\varepsilon \approx (I_4 - I_8)/15$, we get an estimate for the error, and therefore a corrected value for I. So, in this case, we can be reasonably sure that I_8 was correct to within 0.00001, we have every hope that $I_8 - \varepsilon$ is correct to within about 0.000002, and if these results are not good enough, we can expect I_{16} to be correct to within about 0.0000005 and to have an ε that will probably give us 7 decimal places. A lot of gain for trivial pain!

This process is called Richardson Extrapolation, and can be applied to any situation where we know how the error is expected to behave when we change the number of strips [or whatever].

Recommended layout:

Doing numerical work is always much easier if the work is set out logically and neatly. Here is one way to set out a Simpson's Rule calculation [and again we use our standard integral, $I = \int_1^2 1/x \, dx = \log 2 \approx 0.69314718$, working to 8dp]:

$f(1)$	1.00000000		
$f(2)$	0.50000000		
	add:	ends = 1.50000000	
2 strips,	$h = 0.5$		
$f(1.5)$	0.66666667	evens = 0.00000000	
		odds = 0.66666667	
		I_2 = 0.69444444	
4 strips,	$h = 0.25$		
$f(1.25)$	0.80000000		
$f(1.75)$	0.57142857	evens = 0.66666667	
	add:	odds = 1.37142857	$(I_2 - I_4)/15 \approx$
		I_4 = 0.69325397	0.00007936
		subtract:	0.69317461
8 strips,	$h = 0.125$		
$f(1.125)$	0.88888889		
$f(1.375)$	0.72727273		
$f(1.625)$	0.61538462		
$f(1.875)$	0.53333333	evens = 2.03809524	
	add:	odds = 2.76487957	$(I_4 - I_8)/15 \approx$
		I_8 = 0.69315453	0.00000663
		subtract:	0.69314790
16 strips,			
	etc.		

So we can rely on $I \approx 0.69315$ to 5dp, and hope that the final 0.69314790 is a little more accurate than that. [This process is very easy to write a computer program for; it's bad luck if you ever need I_{16} by hand!]

What can go wrong?

Firstly, beware spurious convergence. if we try to evaluate, for example, $\int_0^\pi f(x)\sin(16x)\,\mathrm{d}x$, then, no matter what f looks like, we will get $I_2 = I_4 = I_8 = I_{16} = 0$, which looks very convincing! Only with I_{32} will we finally get some non-zero values.

More serious is what happens if f is not well-behaved. Recall that the error depends on the fourth derivative of f. This will go wrong if f is not continuous [eg $f(x) = 0$ if $x < 0$, 1 if $x \geqslant 0$], or if it has a 'corner' [eg $f(x) = 0$ if $x < 0$, x if $x \geqslant 0$]. It will also go wrong in cases like $f = \sqrt{x}$ near $x = 0$, so that $f'(x)$ is badly behaved. [To get 5sf in $\int_0^1 \sqrt{x}\,\mathrm{d}x$, we need to go to 1024 strips.] Note that even really bad cases will still work *eventually*; but you may have to go to millions of strips. This is not a good idea by computer, and definitely not good by hand.

OK, so $\int_0^1 \sqrt{x}\,\mathrm{d}x = [\frac{2}{3}x^{3/2}]_0^1 = \frac{2}{3}$, and we can do that by hand rather than by Simpson's Rule. What do we do if we need to use numerical analysis?

Consider, for example, $\int_0^1 \sqrt{\sin x}\,\mathrm{d}x$. This 'looks like' $\sqrt{x}$ for x small, so Simpson's Rule will again behave very poorly, and we will need thousands of strips to get decent values. Two ideas.

Firstly, we can use a substitution to get rid of the poor behaviour: if $u^2 = x$, then the integral is $\int_0^1 \sqrt{\sin(u^2)}\,2u\,\mathrm{d}u$, which may not look much better, but now the integrand 'looks like' $2u^2$ for small u, and that's fine [and we get 4dp from 8 strips].

Secondly, we can 'subtract off' the bad behaviour and do some of the work analytically. The integral is $\int_0^1 (\sqrt{\sin x} - \sqrt{x} + \sqrt{x})\,\mathrm{d}x$, and now we can find $\int_0^1 (\sqrt{\sin x} - \sqrt{x})\,\mathrm{d}x$ by Simpson's Rule [which gets an error of less that 0.000001 from 8 strips, even without extrapolation] and add back $\int_0^1 \sqrt{x}\,\mathrm{d}x = \frac{2}{3}$ by hand.

Exercises:

(a) *Straightforward:* Find $\int_0^1 \sin x \, \mathrm{d}x$ numerically to 5dp.

(b) *Needs thought:* Find $\int_0^1 \log \sin x \, \mathrm{d}x \approx -1.05672$ numerically to 4dp. [Logs to base e!] You will have to do something about $x = 0$.

(c) *If time:* Find $\int_0^1 \sqrt{\sin x} \, \mathrm{d}x \approx 0.64298$ to 4dp.