HG3MCE—Computerised Mathematical Techniques in Engineering

Problem Class 1—Solutions

1. The roots are

$$(1234 \pm \sqrt{1234^2 - 4})/2.$$

On my calculator, that is $\frac{1}{2} \cdot (1234 \pm 1233.998379)$, which display as 1233.999190 and 0.0008104. [The 'incorrect' last digit in the smaller root suggests that my calculator is keeping a 'guard' digit.] So the larger root probably has 10sf, and the smaller only 4sf, and we have lost 6sf in finding the smaller root. This is 'cancellation error'—we have subtracted two almost-equal numbers, so that the relative error is hugely magnified.

There are several better ways of getting at the smaller root. Firstly, we know that this quadratic has roots that add to 1234 and multiply to 1 [yes?]. The cancellation error is caused by using addition involving numbers around 1000 to get at an answer around 0.001; we get better results from the product. So the smaller root is 1/1233.999190 [to calculator accuracy].

Secondly, and equivalently, we can 'rationalise the numerator' by multiplying top and bottom by $1234 + \sqrt{1234^2 - 4}$, giving the smaller root as

$$\frac{1}{2} \cdot \frac{1234^2 - (\sqrt{1234^2 - 4})^2}{1234 + \sqrt{1234^2 - 4}} = \frac{1}{2} \cdot \frac{4}{1234 + \sqrt{1234^2 - 4}}$$

and now there is no cancellation error.

Thirdly, we can use a 'process', or 'algorithm'. From the equation, if x is small, we have, ignoring x^2, $-1234x + 1 \approx 0$ or $x \approx 1/1234$. Putting the x^2 back in, we have $-1234x + 1 + x^2 = 0$ [exactly], or $x = (1 + x^2)/1234$. Now, we don't know what x is on the right, but we know it is approximately 1/1234. So a better value for x is $(1 + 1/1234^2)/1234$, which will be a tiny correction [less than one part in a million] to x. If that's not good enough, we can put the corrected value back in to get a further correction [somewhere around the 15th decimal place], and keep going if we need even more accuracy.

By any of these methods, $x \approx 0.0008103733037$ [to 13dp]. Note that since the sum of the roots is 1234, we can get the larger root as 1233.9991896266963 also to 13dp, or 17sf, with almost no extra effort.

You might think that 1234 is a rather large number, and that for 'reasonable' quadratics there is no problem. Firstly, in real life we often get 'large' terms and 'small' terms, and in engineering these can often involve the speed of light or other physical constants, and perhaps exponentials involving such constants. Secondly, even if we replaced 1234 by 12.34, we would be losing a couple of significant figures through cancellation errors. If you allow that, and then use the value of q to calculate something else equally 'carelessly', then after a mere six cycles you've lost 12sf, and your results are worthless, just like the I_{15} we calculated in lectures.

2. Below is the table I obtained on my calculator; yours may be different in detail, but should be similar. Column A is the calculated value of $(\sin(1+h)-\sin 1)/h$ for $h = 10^{-n}$, followed by the number of significant figures of agreement with $\cos 1 \approx 0.540302305$. Column B is similar for $(\sin(1+h)-\sin(1-h))/2h$.

n	A		B	
1	0.497363752	1	0.539402252	2
2	0.536085979	2	0.540293301	4
3	0.53988147	3	0.54030224	6
4	0.5402600	4	0.54030215	6
5	0.540299	5	0.5403025	6
6	0.54026	4	0.540295	5
7	0.5402	3	0.54025	4
8	0.538	2	0.538	2
9	0.49	1	0.53	1

The Taylor series tells us that

$$\sin(1+h) = \sin 1 + h\cos 1 - \tfrac{1}{2}h^2 \sin 1 \ldots$$

and so $(\sin(1+h)-\sin 1)/h \approx \cos 1 - \frac{1}{2}h\sin 1$, and the truncation error is roughly $-\frac{1}{2}h\sin 1$. To confirm this, note for example that when $n = 2$, $h = 0.01$, the error is roughly -0.004.

On the other hand, my calculator is working to 10sf, so the error in working out sines should be no more than $\pm\frac{1}{2}\cdot 10^{-10}$, and the error in the numerator of the formula therefore no more than $\pm 10^{-10}$. So the rounding error in working out the formula is no more than $\pm 10^{-10}/h = \pm 10^{n-10}$. For example, when $n = 6$, the actual error is -0.00004 [nearly all rounding error, as the truncation error is much smaller than this]. The total error is therefore bounded in magnitude by roughly $\frac{1}{2}h\sin 1 + 10^{-10}/h$. This is a minimum when $h^2 = 2\times 10^{-10}/\sin 1$, so $h \approx 1.5\times 10^{-5}$ [just differentiate wrt h and set the derivative to zero].

The formula used for column B has a much smaller truncation error, because of its symmetry. In fact, the truncation error is proportional to h^2, so you need the next term in the Taylor series, and we get a much better result for relatively large values of h [before the cancellation errors get really bad]. Details left as an exercise. Numerical differentiation is always difficult. If you really have to do it, then you need to take h as large as you possibly can, so that you don't get serious cancellation errors, and then you need to use a complicated formula which takes into account ever more terms in the Taylor series so that you don't get truncation errors.