Numerical Methods - Errors (10 pages; 27/3/20)

(1) Reasons why approximations or errors might occur

- simplifying assumptions (eg gravity doesn't vary with height)

- problem with obtaining accurate measurements (technical/cost/effort)

- rounding errors (increasing with the number of calculations); unable to process large numbers of dps

- subtracting nearly equal quantities

- 'ill-conditioned' problems, where a small change in the input value can cause a large change in the output value

(2) Error terminology

Example

Exact value $= 1.512$ (x)

[In practice, the exact value is likely to be an irrational number, rather than a terminating decimal.]

Approximate value $= 1.5$ (X) (eg 1.512 rounded to 1 dp or 2sf)

(Absolute) Error $= X - x = 1.5 - 1.512 = -0.012$

Note: Some textbooks define the Error to be $x - X$

Relative Error $= \frac{X-x}{x} = \frac{-0.012}{1.512} = -0.00794$ (3sf) (provided $x \neq 0$)

Note: Use of the term 'absolute' has traditionally meant taking the modulus of something; so that "the absolute value of x" means $|x|$. With the MEI 2017 specification, however, it has been decided to use "absolute error" to mean "error", as opposed to "relative error". $|X - x|$ and $|\frac{X-x}{x}|$ used to be referred to as the absolute error and absolute relative error, respectively.

Percentage error $= \frac{X-x}{x} \times 100 = 0.794\%$

(3) Obtaining a rounded value from an interval

Example

If $0.9876 < x < 0.9888$, give x to the highest possible number of dps

$\Rightarrow \quad x = 1.0$ to 1dp

$x = 0.99$ to 2dp

but from 0.988 to 0.989 to 3dp

So value of x to highest possible number of dps is 0.99 to 2dp

(4) Obtaining an interval estimate from a rounded value

Example

$\mathrm{X} = 12300$ (3sf) $\Rightarrow 12250 \leq x < 12350$

whilst $\mathrm{X} = 12300$ (4sf) $\Rightarrow 12295 \leq x < 12305$

(5) Errors associated with rounding

(i) Maximum possible error that can occur when rounding to 1dp

Example

eg $X = 1.2 \Rightarrow 1.15 \leq x < 1.25$

$\Rightarrow$ max. possible error = 0.05

(ii) Maximum possible error that can occur when rounding to 1sf:

no limit! (eg $x = 12300000000,\ X = 10000000000$)

(6) Size of relative error

Example

If $2.36 < x < 2.58$, then $X = 2.47$ minimises the error

The relative error will be greatest when the denominator is smallest:

so $\frac{0.11}{2.36} = 0.0466$ (3sf) is the upper limit for the relative error

(7) Accuracy of answers when carrying out arithmetic

(i) **Example**: If $x = 1.2$ to 1 dp and $y = 3.47$ to 2 dp, find an interval estimate for $x + y$

To what accuracy can a single value for $x + y$ be quoted?

Solution

The smallest possible value for $x + y$ is $1.15 + 3.465 = 4.615$

The largest possible value for $x + y$ is $< 1.25 + 3.475 = 4.725$

Thus the interval estimate is $[4.615, 4.725)$

4.615 & 4.725 both round to 5, so value is 5 to 0 dp

(ii) **Example**: If $x = 1.2$ to 1 dp and $y = 3.47$ to 2 dp, find an interval estimate for $y - x$. To what accuracy can a single value for $y - x$ be quoted?

Solution

Smallest possible value for $y - x$ is 3.465 - 1.25 $= 2.215$

Largest possible value for $y - x$ is 3.475 - 1.15 $= 2.325$

Interval estimate is (2.215, 2.325)

2.215, 2.325 both round to 2, so value is 2 to 0 dp

(iii) **Example:** To how many dps is it safe to quote the result of the following addition?

$1.36587 + 1.29166 + 1.32441$

(where each number has been rounded to 5dp)

Solution

Min. $= 1.365865 + 1.291655 + 1.324405 = 3.981925$

Max. $= 1.365875 + 1.291665 + 1.324415 = 3.981955$

5 dp: (3.98193 , 3.98196)

4 dp: (3.9819 , 3.9820)

3 dp: (3.982 , 3.982)

(iv) **Example**: To how many dps is it safe to quote $\frac{4.78256}{2.19982}$?

(where each number has been rounded to 5dp)

Solution

Min. $= \frac{4.782555}{2.199825} = 2.1740616$

Max. $= \frac{4.782565}{2.199815} = 2.1740760$

5 dp: $(2.17406\,,\ 2.17408)$

4 dp: $(2.1741\,,\ 2.1741)$

(8) Propagation of relative errors

(i) Let x be approximated by X and y be approximated by Y

Then relative error iro (in respect of) x is $r_x = \frac{X-x}{x}$ and $r_y = \frac{Y-y}{y}$

Then relative error iro xy is $r_{xy} = \frac{XY-xy}{xy}$

Result to prove: $r_{xy} \approx r_x + r_y$

$r_x + r_y = \frac{X-x}{x} + \frac{Y-y}{y} = \frac{Xy-xy+xY-xy}{xy} = \frac{XY-xy}{xy} + \frac{-XY+Xy+xY-xy}{xy}$

2nd term $= \frac{X(y-Y)+x(Y-y)}{xy} = \frac{(X-x)(y-Y)}{xy} = -r_x r_y$

As $r_x r_y$ is small compared to r_x and r_y, $r_x + r_y \approx r_{xy}$

(ii) **Example**

If $x = 1.414214$ is rounded to $X = 1.414$ and

$y = 1.732051$ is rounded to 1.732, what will be the approximate relative error in xy if it is rounded to XY?

Compare it to the actual value.

Solution

$r_{xy} \approx r_x + r_y = \frac{1.414-1.414214}{1.414214} + \frac{1.732-1.732051}{1.731051}$

$= -0.0001513 - 0.0000295 = -0.0001808$

Actual $r_{xy} = \frac{2.449048-2.449490773}{2.449490773} = -0.0001808\ (4sf)$

(iii) **Exercise**: Find an expression for $r_{\frac{x}{y}}$

Hint: $r_{xy} \approx r_x + r_y \Rightarrow r_{xy} - r_x \approx r_y$

Solution

Let $z = xy$, so that $y = \frac{z}{x}$

Then $r_z - r_x \approx r_{\frac{z}{x}}$

Re-labelling: $r_{\frac{x}{y}} \approx r_x - r_y$

Note: r_y could be -ve

(iv) **Example**: Compare the estimated and actual values of $r_{\frac{x}{y}}$ when $x = 1.414214$ is rounded to $X = 1.414$ and $y = 1.732051$ is rounded to $Y = 1.732$ (as before).

$r_{\frac{x}{y}} \approx r_x - r_y = \frac{1.414-1.414214}{1.414214} - \frac{1.732-1.732051}{1.731051}$

$= -0.0001513 - (-0.0000295) = -0.0001218$

Actual $r_{\frac{x}{y}} = \frac{0.816397229-0.816496743}{0.816496743} = -0.000121879$

(v) Also, it can be shown that $|r_{xy}| \approx |r_x| + |r_y|$

and $|r_{\frac{x}{y}}| \approx |r_x| + |r_y|$

(9) Changing the order of a sequence of operations

$(83 + 55) \times 39 = 5382$

If a computer rounds everything to 2 sf :

(A) $(83 + 55) \times 39$

Step 1: $83 + 55 = 138$, which rounds to 140

Step 2: $140 \times 39 = 5460$, which rounds to 5500

(B) $(83 \times 39) + (55 \times 39)$

Step 1: $83 \times 39 = 3237$, which rounds to 3200

Step 2: $55 \times 39 = 2145$, which rounds to 2100

Step 3: $3200 + 2100 = 5300$

Unrounded: 5382 (A) 5500 (B) 5300

(10) Subtraction involving numbers of a similar size

(i) What number of sig. figs could be quoted for

$467219 - 203426$? (both to 6sf)

Smallest possible value is $467218.5 - 203426.5 = 263792$

Largest possible value is $467219.5 - 203425.5 = 263794$

So 263790 to 5sf

(ii) What number of sig. figs could be quoted for

452.683 - 452.872? (both to 6sf)

Smallest possible value is 452.8715 - 452.6835 = 0.1880

Largest possible value is 452.8725 - 452.6825 = 0.1900

So 0.19 (2dp)

So the number of sig. figs is reduced considerably if the numbers are nearly equal.

(11) 'Ill-conditioned' problems

This is where a small change in the value input results in a large change in the value output

(i) **Example**

If the discriminant of the quadratic equation

$ax^2 + bx + c = 0$ is close to 0,

then a small change in one of a, b or c could result in a change from two solutions to no solutions.

(ii) **Example**

The point of intersection of the lines $y = 2x + 1$ and $y = 2.01x + 9$ will be affected significantly if the gradient of the 2nd line is increased to 2.02

original point is given by:

$2x + 1 = 2.01x + 9 \Rightarrow x = \frac{1-9}{0.01} = -800$

new point is given by:

$2x + 1 = 2.02x + 9 \Rightarrow x = \frac{1-9}{0.02} = -400$

[In such cases, the effects of rounding approximations will be greater; so include extra figures.]

(iii) Error in a function

Let $X = x + h$ be an estimate for x, so that the error is h.

Then the error in $f(x)$ is $f(x+h) - f(x)$.

Now, $f'(x) \approx \frac{f(x+h)-f(x)}{h}$, so that

$f(x+h) - f(x) \approx f'(x)h$

ie the error is magnified by $f'(x)$

(12) Other rounding issues

(i) Premature rounding (using a rounded figure in a calculation, when a more accurate figure could have been used instead)

(ii) The following 'progressive rounding' shouldn't be applied:

$1.23456 \rightarrow 1.2346 \rightarrow 1.235 \rightarrow 1.24$

(iii) 'Chopping': $1.23456 \rightarrow 1.234$ (to be avoided in manual calculations, but may occur with computers)

(13) Exercise

If a measurement of 240 is taken and the % error is revealed to be no more than 5%, find an interval estimate for the exact length.

$X = 240 \;\; \& \;\; 100\left|\frac{X-x}{x}\right| \leq 5$

[as the % error could be negative]

$\Rightarrow -0.05 \leq \frac{240-x}{x} \leq 0.05$

$\Rightarrow (1) \;\; -0.05x \leq 240 - x \;\; (\text{as } x > 0)$

$\Rightarrow \;\; 0.95x \leq 240 \Rightarrow x \leq 252.63(2dp)$

& (2) $240 - x \leq 0.05x$

$\Rightarrow 240 \leq 1.05x$

$\Rightarrow x \geq 228.57(2dp)$

Thus interval estimate is: $228.57 \leq x \leq 252.63$

(14) Exercise

A length is measured as $40.237m$. If the absolute size of the relative error is no more than 0.05, what is the greatest possible absolute size of the absolute error, to 2 dp?

Solution

Let x be the true value.

Then $\left|\frac{40.237-x}{x}\right| \leq 0.05$, so that $-0.05 \leq \frac{40.237-x}{x} \leq 0.05$

$\Rightarrow x(1-0.05) \leq 40.237$ and $40.237 \leq x(1+0.05)$ (as $x > 0$)

$\Rightarrow x \leq 42.35474$ and $x \geq 38.32095$ (5dp)

If $x = 38.32095$, then absolute error is $40.237 - 38.32095 = 1.91605$

If $x = 42.35474$, then absolute error is $40.237 - 42.35474 = -2.11774$

So the greatest possible absolute size of the absolute error is 2.11774 or 2.12 (2dp).