PROBLEM SET #1

APC 523/MAE 507/AST 523 : Numerical Algorithms for Scientific Computing

Vivek Kumar

March 13, 2019

1 Error in (symmetric) rounding vs chopping

Assertion: When mapping a real number x to a nearby machine number in $\mathbb{R}(p,q)$, the upper bound in the relative error for symmetric rounding is:

$$\left|\frac{x - \mathrm{rd}(x)}{x}\right| \leq 2^{-p}$$

Proof:
Consider the number x to be represented as:

$$x = \pm\left(\sum_{l=1}^{\infty} b_{-l}2^{-l}\right)2^e$$

If the number is to be rounded to p terms, two cases arise:

CASE I. The $(p+1)^{\text{th}}$ is 0.
In this scenario the difference between the true value and the rounded value is given by:

$$x - \mathrm{rd}(x) = \pm\left(\sum_{l=p+2}^{\infty} b_{-l}2^{-l}\right)2^e$$

The maximum relative error can then be computed as:

$$\begin{aligned}
\max\left|\frac{x - \mathrm{rd}(x)}{x}\right| &= \frac{\max|x - \mathrm{rd}(x)|}{\min|x|} \\
&= \frac{2^{-p-1}2^e}{2^{-1}2^e} \\
&= 2^{-p}
\end{aligned}$$

which is what we set to prove.

CASE II. The $(p+1)^{\text{th}}$ is 1.
In this scenario the maximum difference between the true and the rounded value is obtained as:

$$\max|x - \mathrm{rd}(x)| = \left(2^{-p} - 2^{-p-1}\right)2^e$$

This is the case we all the leading terms from $(p+2)$ are 1. Hence the maximum relative error can be computed as before:

$$\begin{aligned}
\max\left|\frac{x - \mathrm{rd}(x)}{x}\right| &= \frac{\max|x - \mathrm{rd}(x)|}{\min|x|} \\
&= \frac{\left(2^{-p} - 2^{-p-1}\right)2^e}{2^{-1}2^e} \\
&= 2^{-p}
\end{aligned}$$

which is what we set to prove

Both the cases show that the maximum symmetric rounding off error is 2^{-p}

2 Recurrence in reverse

(a) The reverse recurrence relation is given by:

$$y_{n-1} = \frac{e - y_n}{n}$$

Computing for a few terms down the chain we obtain:

$$\begin{aligned} y_{n-2} &= \frac{e - y_{n-1}}{n-1} \\ &= \frac{ne - e + y_n}{n\left(n-1\right)} \\ y_{n-3} &= \frac{e - y_{n-2}}{n-2} \\ &= \frac{n\left(n-1\right)e - ne + e - y_n}{n\left(n-1\right)\left(n-2\right)} \end{aligned}$$

One can denote the pattern as:

$$y_{n-p} = (-1)^p \frac{y_n}{n\left(n-1\right)\left(n-2\right)\ldots\left(n-p+1\right)} + e\left[\frac{1}{n-(p-1)} - \frac{1}{\left(n-(p-1)\right)\left(n-(p-2)\right)} + \frac{1}{\left(n-(p-1)\right)\left(n-(p-2)\right)\left(n-(p-3)\right)} + \ldots\right]$$

To obtain the value of y_k in terms of y_N we replace $n-p$ with k and simplify:

$$\begin{aligned} y_k &= (-1)^{n-k}\frac{y_n}{n\left(n-1\right)\left(n-2\right)\ldots\left(k+1\right)} + \text{exponent terms} \\ &= (-1)^{n-k}\frac{y_n k!}{n!} + \text{exponent terms} \end{aligned}$$

The condition number is given as:

$$\begin{aligned} \left(\text{cond } g_k\right)\left(y_k\right) &= \left|\frac{y_N g'\left(y_N\right)}{y_k}\right| \\ &= \left|\frac{y_N \frac{k!}{N!}}{y_k}\right| \end{aligned}$$

Since the k is less than N, y_k is greater than y_N, the upper bound on the condition number, $\left(\text{cond } g_k\right)\left(y_k\right)$, obtained as:

$$\left(\text{cond } g_k\right)\left(y_k\right) \leq \frac{k!}{N!}$$

as $\frac{y_N}{y_k} \leq 1$

(b) We know the condition number represents:

$$\begin{aligned} \varepsilon_y &= \left(\text{cond } g_k\right)\varepsilon_x \\ \frac{\Delta y_k}{y_k} &\leq \frac{k!}{N!} \quad \leq \varepsilon \\ N! &\geq \frac{k!}{\varepsilon} \end{aligned}$$

Here, we have assumed $\varepsilon_x = 1$ and ε is a predefined target error in y_k.

(c) For `python3` the machine epsilon for float is $1.0e^{-15}$(Obtained using numpy.finfo(float)). Using this machine epsilon the value of N obtained is 31. [Check code]

(d) The computed value of y_{20} is 0.123803830762570 and the value of y_{20} directly by integration is 0.123803830762570. [Check code]