Dihan Yang **CS 450** **HW#1**

Q1

$$g(x,h)=\frac{-f(x+2h)+8f(x+h)-8f(x-h)+f(x-2h)}{12h}$$

1. What derivative is approximated in the above formula?

   The formula is approximating a **first-order derivative**, because in general, this function is calculating the ratio between the difference of value and difference of parameter.

2. What is the leading order truncation error?

   Use Taylor Series to expand 4 terms in $g(x,h)$:

$$f(x+h)=f(x)+\frac{h^1}{1!}f'(x)+\frac{h^2}{2!}f''(x)+\frac{h^3}{3!}f'''(x)+\frac{h^4}{4!}f''''(x)+\cdots$$

$$f(x+2h)=f(x)+\frac{(2h)^1}{1!}f'(x)+\frac{(2h)^2}{2!}f''(x)+\frac{(2h)^3}{3!}f'''(x)+\frac{(2h)^4}{4!}f''''(x)+\cdots$$

$$f(x-h)=f(x)+\frac{(-h)^1}{1!}f'(x)+\frac{(-h)^2}{2!}f''(x)+\frac{(-h)^3}{3!}f'''(x)+\frac{(-h)^4}{4!}f''''(x)+\cdots$$

$$f(x-2h)=f(x)+\frac{(-2h)^1}{1!}f'(x)+\frac{(-2h)^2}{2!}f''(x)+\frac{(-2h)^3}{3!}f'''(x)+\frac{(-2h)^4}{4!}f''''(x)+\cdots$$

   Plug into $g(x,h)$:

$$\mathrm{g(x,h)}=\frac{8\big(f(x+h)-f(x-h)\big)+\big(f(x-2h)-f(x+2h)\big)}{12h}$$

$$=\frac{8\left(\frac{2h}{1!}f'(x)+\frac{2h^3}{3!}f'''(x)+\frac{2h^5}{5!}f'''''(x)+\cdots\right)-\left(\frac{2(2h)}{1!}f'(x)+\frac{2(2h)^3}{3!}f'''(x)+\frac{2(2h)^5}{5!}f'''''(x)+\cdots\right)}{12h}$$

$$=\frac{8\left(h\cdot f'(x)+\frac{h^3}{3!}f'''(x)+\frac{h^5}{5!}f'''''(x)+\cdots\right)-\left(2h\cdot f'(x)+\frac{2^3h^3}{3!}f'''(x)+\frac{2^5h^5}{5!}f'''''(x)\right)}{6h}$$

$$=\frac{8h\cdot f'(x)-2h\cdot f'(x)}{6h}+\frac{\frac{8h^3-8h^3}{3!}f'''(x)}{6h}+\frac{\frac{8h^5-32h^5}{5!}f'''''(x)}{6h}+\cdots$$

$$=f'(x)+0-\frac{4h^4}{5!}f'''''(x)+\cdots$$

The first term $f'(x)$ is the exact term, while the rest are truncation errors. Within the truncation error, h^1, h^2, h^3 term all get cancelled out, and

the existed leading order absolute truncation error is $|\frac{h^4}{30} f^{(5)}(x)|$

3. What is the estimated round-off error as a function of h? (Assume condition number is roughly 1).

Let $f_1 = f(\hat{x}+\hat{h})$, $f_2 = f(\hat{x}-\hat{h})$, $f_3 = f(\hat{x}+2\hat{h})$, $f_4 = f(\hat{x}-2\hat{h})$

$$\mathrm{fl}\left(\frac{\delta f}{\delta x}\right) = \frac{8(\widehat{f_1} - \widehat{f_2}) + (\widehat{f_4} - \widehat{f_3})}{12h}$$

$$\because \text{condition number} = \frac{\Delta f/f}{\Delta x/x} \cong 1$$

$$\frac{\Delta x}{x} \le \epsilon_{mach}$$

$$\therefore \frac{\Delta f}{f} = \frac{\hat{f} - f}{f} \le \epsilon_{mach}$$

$$\therefore \hat{f} \le f(1 + \epsilon_{mach})$$

Let $\widehat{f_1} = f_1(1+\epsilon_1)$, $\widehat{f_2} = f_2(1+\epsilon_2)$, $\widehat{f_3} = f_3(1+\epsilon_3)$, $\widehat{f_4} = f_4(1+\epsilon_4)$

Where $\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4 \le \epsilon_{\mathrm{mach}}$

$$\mathrm{fl}\left(\frac{\delta f}{\delta x}\right) \cong \frac{8(f_1 + f_1\epsilon_1 - f_2 - f_2\epsilon_2) + (f_4 + f_4\epsilon_4 - f_3 - f_3\epsilon_3)}{12h}$$

$$\cong \frac{8(f_1 - f_2) + (f_4 - f_3)}{12h} + \frac{8(f_1\epsilon_1 - f_2\epsilon_2) + (f_4\epsilon_4 - f_3\epsilon_3)}{12h}$$

$$\cong \frac{8(f_1 - f_2) + (f_4 - f_3)}{12h} + \frac{8(\epsilon_1 - \epsilon_2) + (\epsilon_4 - \epsilon_3)}{12h} f(x)$$

$$\frac{8(\epsilon_1 - \epsilon_2) + (\epsilon_4 - \epsilon_3)}{12h} f(x) \le \frac{18\epsilon_{mach}}{12h} |f(x)| = \frac{3\epsilon_{mach}}{2h} |f(x)|$$

Therefore, the **absolute round-off error will be** $\boldsymbol{\frac{3\epsilon_{mach}}{2h} |f(x)|}$,

where as the bound for **relative round-off error will be** $\boldsymbol{\frac{3\epsilon_{mach}}{2h}}$

4. For what h is the minimal total error attained?

   Total Error = Truncation Error + Round-off Error

$$\mathrm{E} = \frac{3\epsilon_{\mathrm{mach}}}{2h}|f(x)| + \frac{h^4}{30}|f^{(5)}(x)|$$

   In order to achieve minimum E, set the first derivative=0

$$\frac{dE}{dh} = \frac{-3\epsilon_{mach}}{2h^2}|f(x)| + \frac{4h^3}{30}|f^{(5)}(x)| = 0$$

$$\therefore \mathbf{h} = \sqrt[5]{\left|\frac{45\epsilon_{mach}f(x)}{4f^{(5)}(x)}\right|} \textbf{ for minimal total error.}$$

5. What is the estimated minimal *total error* (for IEEE 64-bit arithmetic)?

   Plug the h obtained in last question into the expression of total error, with $\epsilon_{\mathrm{mach}} = 2^{-53} \cong 1.11\times10^{16}$ for 64-bit arithmetic.

$$\mathrm{E} = \frac{3\epsilon_{mach}}{2h}|f(x)| + \frac{h^4}{30}f^{(5)}(x)$$

$$= \frac{3\epsilon_{mach}}{2\left(\left|\frac{45\epsilon_{mach}f(x)}{4f^{(5)}(x)}\right|\right)^{\frac{1}{5}}}|f(x)| + \frac{\left(\left|\frac{45\epsilon_{mach}f(x)}{4f^{(5)}(x)}\right|\right)^{\frac{4}{5}}}{30}f^{(5)}(x)$$

$$= \epsilon_{mach}^{\frac{4}{5}} \cdot g\left(f^{\frac{4}{5}}(x), f^{(5)\frac{1}{5}}(x)\right)$$

$$\therefore \textbf{the estimated minimal } \boldsymbol{total\ error} \textbf{ would be the order of } \boldsymbol{\epsilon_{mach}^{\frac{4}{5}}}$$