DERIVATIVES

i) Definition $\frac{df}{dx} \equiv f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$

Simplest thing to do. Take h "small" and estimate $f'(x)$ via finite difference

$$\frac{df}{dx} \simeq \frac{f(x+h) - f(x)}{h} \quad ; \; h > 0, \; h \text{ small.} \Rightarrow \text{forward difference}$$

OR

$$\frac{df}{dh} \simeq \frac{f(x) - f(x-h)}{h} \quad ; \; h > 0, \; h \text{ small} \Rightarrow \text{backward difference}$$

ii) Problem: with difference, when h gets very small, you get large rounding errors.

•) Let's estimate the numerical error. Assume machine precision $C \simeq 10^{-16}$

•) As usual, Taylor expand around x:

$$f(x+h) = f(x) + \frac{df}{dx} h + \frac{1}{2} \frac{d^2f}{dx^2} h^2 + O(h^3)$$

$$\Downarrow$$

$$\frac{df}{dx} = \underbrace{\frac{f(x+h) - f(x)}{h}}_{\text{Our numerical estimate}} - \underbrace{\frac{1}{2} h \frac{d^2f}{dh^2}}_{\text{Numerical error at leading order}} + \cdots$$

•) Therefore $|\varepsilon_1| = \left|\frac{f''(x)}{2}\right| h$ is the numerical error introduced by truncating the expansion.

•) But there is a rounding error in $f(x+h) - f(x)$ as well. We know that, if we evaluate $f(x)$, we have a rounding error on $f(x)$ which is $\varepsilon_r \simeq C f(x)$

Since $f(x+h) \simeq f(x)$ we get:

$$\begin{cases} \varepsilon_{r, f(x+h)} \simeq C f(x+h) \\ \varepsilon_{r, f(x)} \simeq C f(x) \end{cases}$$

Therefore the error on the difference is going to be $\lesssim 2C f(x)$ and the overall rounding error is then

$$|\varepsilon_2| \lesssim \frac{2C\,|f(x)|}{h}$$

Our conservative assumption on the total error is thus

$$\boxed{\varepsilon_{der} = |\varepsilon_1| + |\varepsilon_2| = \frac{h}{2}\left|f''(x)\right| + \frac{2C\,|f(x)|}{h}} \quad (*)$$

•) Note that, if we decrease h, $|\varepsilon_1|$ decreases but $|\varepsilon_2|$ increases at the same time! It is thus not true that arbitrarily small h minimizes the error

To minimize ε_{der} let us compute $\boxed{\frac{d\varepsilon_{der}}{dh} = 0} \Rightarrow \frac{|f''|}{2} - \frac{2C}{h^2}|f| = 0$

Therefore $\mathcal{E}_{der}^{MIN}$ is obtained for:

$$\boxed{h = \left(4C \frac{|f|}{|f''|}\right)^{\frac{1}{2}}} \; ; \quad \text{this is the optimal choice} \qquad (**)$$

Assuming $|f(x)| \sim 1$ and $|f''(x)| \simeq 1$ (not always true!), we get ($C \simeq 10^{-16}$) an optimal choice of $\boxed{h \sim 10^{-8}}$

•) Replacing $(**)$ into $(*)$ we find a minimum error:

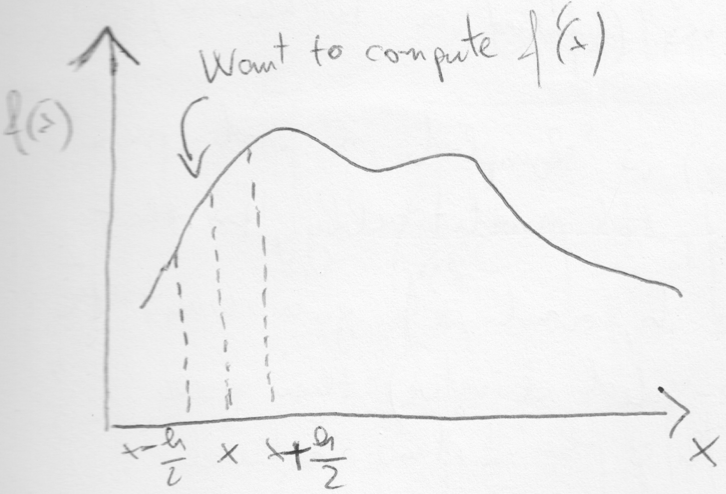
$$\mathcal{E}_{der} \simeq \frac{\cancel{2}\sqrt{C}}{\cancel{2}} |f''|^{\frac{1}{2}} |f|^{\frac{1}{2}} + \frac{\cancel{2}C}{\cancel{2}C^{\frac{1}{2}}} \frac{|f|}{|f|^{\frac{1}{2}}} |f''|^{\frac{1}{2}} \Rightarrow$$

$$\boxed{\mathcal{E}_{der} = \left[C \, |f(x)| \, |f''(x)|\right]^{\frac{1}{2}}}$$

If $|f| \sim 1$ and $|f''| \sim 1$, we are always limited to $\mathcal{E} \sim 10^{-8}$, that is 8-digits precision in the derivative.

iii) To improve precision, use central differences

$$\boxed{\frac{df}{dx} \simeq \frac{f\left(x + \frac{h}{2}\right) - f\left(x - \frac{h}{2}\right)}{h}}$$



•) You can show as exercise that $f'(x) = \frac{f(x+\frac{h}{2}) - f(x-\frac{h}{2})}{h} - \frac{1}{24}h^2 f'''(x) + \dots$

(just Taylor expand both $f(x+\frac{h}{2})$, $f(x-\frac{h}{2})$ and rearrange terms suitably)

Therefore now ε_1 is $O(h^2)$ (before it was $O(h)$).

We still have same rounding error of course. Hence

$$\boxed{|\varepsilon|^{\text{central der.}} = \frac{2C|f|}{h} + \frac{h^2}{24}|f'''(x)|}$$

•) The optimal h becomes

$$\boxed{h = \left(24\, C \left|\frac{f(x)}{f'''(x)}\right|\right)^{\frac{1}{3}}}$$

If $f(x) \sim 1$, $f'''(x) \sim 1$ and $C \simeq 10^{-16}$, we now get

$$\boxed{h \simeq 10^{-5}}$$

•) The corresponding error is

$$\boxed{\varepsilon = \left(\frac{9}{8} C^2 |f(x)|^2 |f'''(x)|\right)^{\frac{1}{3}}}$$

For $f \sim 1$, $f''' \sim 1$, $C \simeq 10^{-16}$ we get $\varepsilon \sim 10^{-10}$. About 100 times better than the 1-side derivatives (forward or backward)

•) Note that, very often, your $f(x)$ is sampled in fixed points from data and cannot be computed analitically. In that case you don't get to choose h, and, if you want to take the central difference (2-sided derivative) then your interval will be $2h$, where h is the distance between consecutive points in the sample

iv) We often need to compute second derivatives: $f''(x)$. Of course, we do that by finite difference, ~~and~~ starting from first derivative. We use central, 2-sided scheme for everything:

$$\begin{cases} f''(x) \simeq \dfrac{f'\left(x+\frac{h}{2}\right) - f'\left(x-\frac{h}{2}\right)}{h} \\[2ex] f'\left(x+\frac{h}{2}\right) \simeq \dfrac{f(x+h)-f(x)}{h} \quad ; \quad f'\left(x-\frac{h}{2}\right) = \dfrac{-f(x-h)+f(x)}{h} \end{cases}$$

Combining everything:

$$f''(x) \simeq \frac{1}{h}\left[\frac{f(x+h)-f(x)}{h} - \frac{f(x)-f(x-h)}{h}\right] =$$

$$= \frac{1}{h}\,\frac{f(x+h)-2f(x)+f(x-h)}{h}$$

Therefore: $\boxed{f''(x) \simeq \dfrac{f(x+h)-2f(x)+f(x-h)}{h^2}}$

.) With the usual Taylor expansions (for $f(x+h)$ and $f(x-h)$) we can estimate ε and the optimal choice of h.

$$\varepsilon = \frac{4C|f|}{h^2} + \frac{h^2}{12}|f'''|$$

$$h = \left(48C\left|\frac{f}{f'''}\right|\right)^{\frac{1}{4}} \qquad (h \simeq 10^{-4})$$

$$\varepsilon_{MIN} = \left(\frac{4}{3}C|f||f'''|\right)^{\frac{1}{2}} \qquad (\varepsilon_{MIN} \simeq 10^{-8})$$

We get similar accuracy as 1-sided 1^{st} derivative evaluation.