Problem Set 1

September 17^{th}, 2021
PHYS-512

Jules Faucher
260926201

1. a) First, we need to use Taylor expansions of $f(x\pm\delta)$ and $f(x\pm 2\delta)$.

$$f(x\pm\delta) = f(x) \pm f'(x)\delta + \frac{1}{2}f''(x)\delta^2 \pm \frac{1}{6}f'''(x)\delta^3 + \frac{1}{24}f^{(4)}(x)\delta^4 \pm \frac{1}{120}f^{(5)}(x)\delta^5 + \ldots$$

$$f(x\pm 2\delta) = f(x) \pm 2f'(x)\delta + 2f''(x)\delta^2 \pm \frac{4}{3}f'''(x)\delta^3 + \frac{2}{3}f^{(4)}(x)\delta^4 \pm \frac{4}{15}f^{(5)}(x)\delta^5 + \ldots$$

We can reuse these definitions in the numerical derivatives.

For simplicity, let $\tilde{f}_1' = \dfrac{f(x+\delta) - f(x-\delta)}{2\delta}$

and $\tilde{f}_2' = \dfrac{f(x+2\delta) - f(x-2\delta)}{4\delta}$

So, $\tilde{f}_1' = f'(x) + \frac{1}{6}f'''(x)\delta^2 + \frac{1}{120}f^{(5)}(x)\delta^4 + \ldots$

and, $\tilde{f}_2' = f'(x) + \frac{2}{3}f'''(x)\delta^2 + \frac{2}{15}f^{(5)}(x)\delta^4 + \ldots$

We then need to find a combination of $\tilde{f}_1'$ and $\tilde{f}_2'$ such that the third derivative should be cancelled.

$$\alpha\tilde{f}_1' + \beta\tilde{f}_2' = \alpha f'(x) + \frac{\alpha}{6}f'''(x)\delta^2 + \frac{\alpha}{120}f^{(5)}(x)\delta^4 + \beta f'(x) + \frac{2\beta}{3}f'''(x)\delta^2 + \frac{2\beta}{15}f^{(5)}(x)\delta^4 + \ldots = f'(x) + \mathcal{O}(\delta^4)$$

$$\Rightarrow \frac{\alpha}{6} + \frac{2\beta}{3} = 0 \rightarrow \alpha = -4\beta$$

and $\alpha + \beta = 1$

$$\left.\begin{matrix} \\ \\ \end{matrix}\right\} \quad \alpha = 4/3, \quad \beta = -1/3$$

$$\frac{4}{3}\tilde{f}_1'(x) - \frac{1}{3}\tilde{f}_2'(x) = \frac{4}{3}\left(\frac{f(x+\delta)-f(x-\delta)}{2\delta}\right) - \frac{1}{3}\left(\frac{f(x+2\delta)-f(x-2\delta)}{4\delta}\right) + \ldots$$

$$\therefore f'(x) = \frac{f(x-2\delta)+8f(x+\delta)-8f(x-\delta)-f(x+2\delta)}{12\delta} + \sigma(\delta^5)$$

b) We need to use the truncation error as well as the round-off error.

$$\text{Truncation error} = e_t \approx \left|\frac{4}{3}\tilde{f}_1' - \frac{1}{3}\tilde{f}_2' - f'(x)\right|$$

$$\approx \left|\frac{4}{3}\cdot\frac{1}{120} f^{(5)}(x)\,\delta^4 - \frac{1}{3}\frac{2}{15} f^{(5)}(x)\,\delta^4 + \ldots\right|$$

$$\approx \left|\frac{1}{90} f^{(5)}(x)\,\delta^4 - \frac{2}{45} f^{(5)}(x)\,\delta^4\right|$$

$$\approx \frac{1}{30}\left|f^{(5)}(x)\right|\delta^4$$

$$\text{round-off error} = e_r \approx \varepsilon_m \left|f(x)/\delta\right|$$

We need to find a δ that minimizes $e_r^2 + e_t^2$

$$e_t^2 + e_r^2 = \frac{1}{900}\left(f^{(5)}(x)\right)^2 \delta^8 + \varepsilon_m^2 \frac{f^2(x)}{\delta^2}$$

$$\frac{d(e_t^2+e_r^2)}{d\delta} = \frac{8}{900}\left(f^{(5)}(x)\right)^2 \delta^7 - 2\varepsilon_m^2 \frac{f^2(x)}{\delta^3} = 0$$

$$\Rightarrow \frac{8}{900}\left(f^{(5)}(x)\right)^2 \delta^7 = \frac{2\varepsilon_m^2 f^2(x)}{\delta^3}$$

$$\Rightarrow \delta = \sqrt[5]{225\left(\frac{\varepsilon_m f(x)}{f^{(5)}(x)}\right)^2}$$

To test our step size with e^x, we know that $f^{(5)}(x) = f(x)$

$$\therefore \delta = \sqrt[5]{225\,\varepsilon_m^2} \approx 1{,}6 \times 10^{-6}$$

With $e^{0{,}01}$, $f^{(5)}(x) = (0{,}01)^5 f(x)$

$$\therefore \delta = \sqrt[5]{225 \cdot (0{,}01)^5\,\varepsilon_m^2} = 1{,}6 \times 10^{-8}$$

As we see when running the code, the fractionnal errors on the derivatives are very low. The step size are thus appropriate.

2. We need to evaluate the size of the steps. In class, we showed that $dx \sim \sqrt[3]{\frac{\varepsilon_f f}{f'''}}$. This means we need an expression for $f'''(x)$

We can reuse our expressions from Problem 1, and cancel the $f'(x)$'s

$$f(x \pm \delta) = f(x) \pm f'(x)\delta + \tfrac{1}{2} f''(x)\delta^2 \pm \tfrac{1}{6} f'''(x)\delta^3 + \ldots$$

$$f(x \pm 2\delta) = f(x) \pm 2f'(x)\delta + 2f''(x)\delta^2 \pm \tfrac{4}{3} f'''(x)\delta^3 + \ldots$$

$$f(x+2\delta) - f(x-2\delta) = 4f'(x)\delta + \tfrac{8}{3} f'''(x)\delta^3 + \ldots$$

$$f(x+\delta) - f(x-\delta) = 2f'(x)\delta + \tfrac{1}{3} f'''(x)\delta^3 + \ldots$$

$$\Rightarrow f(x+2\delta) - f(x-2\delta) + 2f(x-\delta) - 2f(x+\delta) = 2f'''(x)\delta^3 + \ldots$$

$$\therefore f'''(x) = \frac{f(x+2\delta) - f(x-2\delta) + 2f(x-\delta) - 2f(x+\delta)}{2\delta^3}$$

3. To create an interpolation, a cubic spline was used. Since cubic polynomials are used, the error should be of forth order ($O(h^4)$ where h is the distance between two points)

As such, the error should be taken to be h_{max}^4 where h_{max} is the largest distance between two data points.

From the deBoor's book "A Practical Guide to Splines", on page 55, eqn 12, we can see that

$$\underbrace{| f(x) - CS(x) |}_{\approx \text{error}} \le \frac{1}{16} |h|^4 | f^{(4)}(x) |$$

We will need the forth derivative at the point where we interpolate the temperature.

Using a similar derivation as seen in problem 4, we can show that

$$f^{(4)}(x) \approx \frac{f(x+2h) - 4f(x+h) + 6f(x) - 4f(x-h) + f(x-2h)}{h^4}$$

We can adapt this formula to our needs, where h depends on the available data.

$$T^{4}(V) = \frac{T[i+2] - 4T[i+1] + 6T[i] - 4T[i-1] + T[i-2]}{\min(V_i - V_j)}$$

4. To compare the accuracies of the different methods, I wrote a routine which plots the differences between the interpolations and the actual functions.

For cos(x), I used 7 points. The best interpolation method is the polynomial interpolation.

For the Lorentzian, I also used 7 point. The original rational function is by far the worse even though the Lorentzian is a rational function itself. Given this, the error should have been close to zero.

We obtain this desired behaviour is obtained when np.linalg.inv() by np.linalg.pinv(). This yield errors on the scale of the machine's precision, which is ideal.

If we take a look at p and q when we used the regular inv function, we can see that $q = 3x + x^2 - x^3$. We can presume that some non-zero matrix element has had its value increased during the inversion, and made their way into our polynomial.

This is why using Pinv fixes the issue. The non-zero matrix elements get removed and do not end up in our polynomials.