AERO-222: Introduction to Aerospace Computation, Fall 2021
Homework #1, Due Date: Wednesday, September 22, 2021

Show all work and justify your answers!

Instructions

- *This homework contains both handwritten and coding problems and shall be submitted according to the following guidelines.*
- *Hardcopy:*
  - *Due on CANVAS at 11:59 PM on the day of the deadline.*
  - *Shall include screenshots of any hand-written work.*
  - *For coding problems, the hardcopy shall include any relevant derivations and emphasize the final results (i.e. boxed, highlighted, etc.).*
  - *Shall be submitted as a single file according to the provided template with the following naming scheme:* "LastnameHW#.pdf"
- *Coding Submission:*
  - *Due on CANVAS at 11:59 PM on the day of the deadline.*
  - *Shall be submitted as a single file according to the provided template with the following naming scheme:* "LastnameHW#.py"
  - *The script shall print out all outputs asked for in the problem.*
- *Late submissions will be accepted with a 10 point deduction per day late.*

---

1. Round-off and Truncation Error (15 pts) Code. Calculate the roots (to an absolute error $\varepsilon_x < 0.001$) of

$$f(x) = x^2 - 4x - 6 = 0$$

Use the following 3 methods:

- Bisection method
- Secant method
- Regula-Falsi method

These methods require a set of starting points/bounds. Use $x_1 = -4$ and $x_2 = +4$, which will meet the initial requirements for all 3 methods. For each method, calculate and plot $f(x_n)$ as a function of the iteration number n. Which method converges the fastest?

Solution:
The Secant Method converges the fastest. Try and run the code. There is a chance they changed the code for each figure, but it should generate a correct solution.

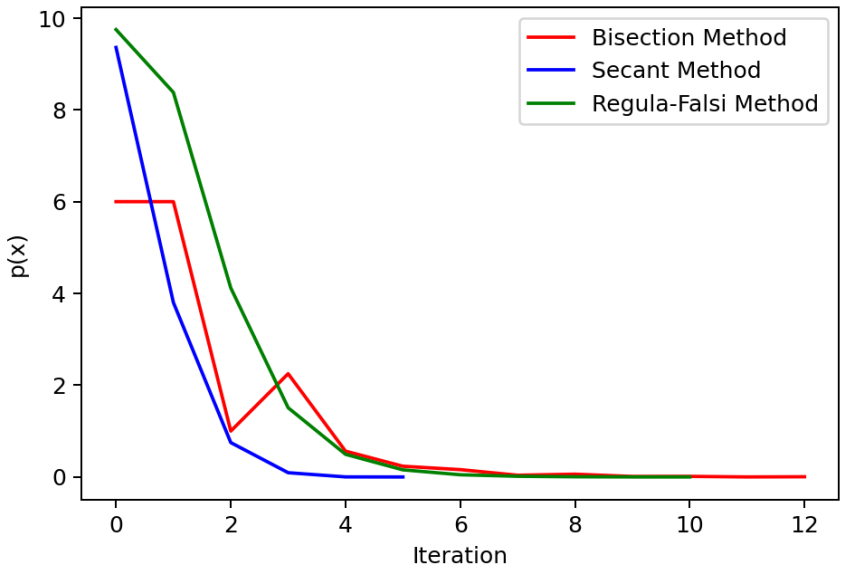

Figure 1: The solution should look very similar to this, all methods should converge to zero

2. Round-off Error (10 pts) Code. Write a script to:

(a) Evaluate the cubic polynomial, $f(x)$, at $x = 1.35$, using default machine precision and again using only three significant digits at each arithmetic operation. Calculate the absolute and relative error of the final result:

$$f(x) = 2.32\,x^3 + 2.08\,x^2 - 4.86\,x + 8.33$$

(b) Repeat part (a) but do it with *nested multiplication*. Compare errors. Which takes fewer operations? The nested form is:

$$f(x) = [(2.32\,x + 2.08)\,x - 4.86]\,x + 8.33$$

Solution:

(a) With default machine precision:

$$2.32 \times 1.35^3 + 2.08 \times 1.35^2 - 4.86 \times 1.35 + 8.33 = 11.2679$$

With rounding to 3 significant digits in between operations (powers are treated as 1 operation):

$$2.32 \times 1.35^3 + 2.08 \times 1.35^2 - 4.86 \times 1.35 + 8.33 = 11.268$$
$$\text{Absolute Error} = 0.130 \times 10^{-3}$$
$$\text{Relative Error} = 0.115 \times 10^{-4}$$
$$\text{Number of Steps} = 9$$

(b) With default machine precision:

$$[(2.32\,x + 2.08)\,x - 4.86]\,x + 8.33 = 11.2679$$

With rounding to 3 significant digits in between operations:

$$[(2.32\,x + 2.08)\,x - 4.86]\,x + 8.33 = 11.268$$
$$\text{Absolute Error} = 0.13 \times 10^{-3}$$
$$\text{Relative Error} = 0.115 \times 10^{-4}$$
$$\text{Number of Steps} = 6$$

In this case, the errors are the same because both results round to the same value for 3 significant digits. The first approach is in general more accurate but requires more operations. If any very small values show different results (smaller than 10^{-10}), its due to machine/python round-off errors.

3. **Truncation Error (20 pts) Code.** Evaluate $f(x) = e^{-4}$ to four digits of precision (chopping) using the following two approaches:

(a) $e^{-4} = \sum_{k=0}^{7} \frac{(-4)^k}{k!} = \sum_{k=0}^{7} \frac{(-1)^k 4^k}{k!}$

(b) $e^{-4} = \frac{1}{e^4} \approx \frac{1}{\sum_{k=0}^{7} \frac{(4)^k}{k!}}$

Note that the true value to four digits of precision is 1.8316×10^{-2}. Which formula gives more accurate results and why? Plot the error of each approach as a function of iteration number.

Solution:

(a) $1 - 4 + 8 - 10.67 + 10.67 - 8.53 + 5.69 - 3.25 = -1.095$

(b) $1/(1 - 4 + 8 - 10.67 + 10.67 - 8.53 + 5.69 - 3.25) = 1.930 \times 10^{-2}$

The actual value is 1.8316×10^{-2} to 4 digits of precision. The second method is more accurate because it does not oscillate about the value, but instead always stays smaller than 1 and larger than 0. Thus, the truncation error of the second method is lower for the same number of steps.

4. **Taylor Series (15 pts) By-hand.** Expand the function, $f(x) = x^3 + \cos x$, by Taylor series up to degree 3 for the following cases:

(a) centered at $x_0 = 4$ and evaluated at $x_1 = x_0 - 0.6$;

(b) centered at $x_0 = 3$ and evaluated at $x_1 = x_0 + 0.1$.

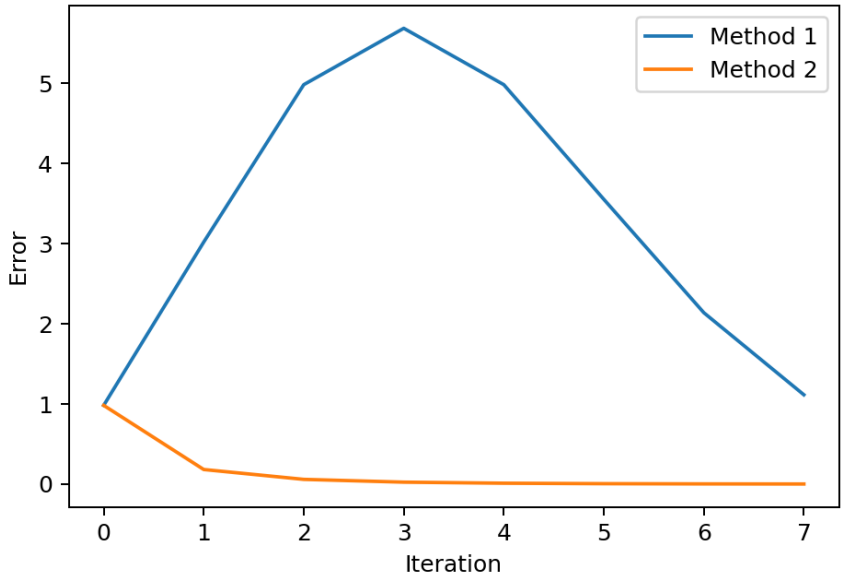

Figure 2: Method 2 is more accurate

Solution:
First write the general form of the Taylor polynomial for clarity:

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)(x - x_0)^2}{2!} + \frac{f'''(x_0)(x - x_0)^3}{3!}$$

Then, insert the given equation:

$$f(x) \approx x_0^3 + \cos x_0 + (3x_0^2 - \sin x_0)(x - x_0) + \frac{(6x_0 - \cos x_0)(x - x_0)^2}{2!} + \frac{(6 + \sin x_0)(x - x_0)^3}{3!}$$

Insert given values:

a. centered $x_0 = 4$ and evaluated $x_0 - 0.6 = 3.4$

$$f(x) \approx 4^3 + \cos 4 + (3(4)^2 - \sin 4)(3.4 - 4) + \frac{(6(4) - \cos 4)(3.4 - 4)^2}{2!} + \frac{(6 + \sin 4)(3.4 - 4)^3}{3!}$$

answer: 38.34

b. centered $x_0 = 3$ and evaluated $x_0 + 0.1 = 3.1$

$$f(x) \approx 3^3 + \cos 3 + (3(3)^2 - \sin 3)(3.1 - 3) + \frac{(6(3) - \cos 3)(3.1 - 3)^2}{2!} + \frac{(6 + \sin 3)(3.1 - 3)^3}{3!}$$

answer: 28.79

5. Variables and Computer Precision (10 pts) By-hand. Answer the following questions:

(a) Which can store a larger number, a signed int or unsigned int?

(b) Which uses more memory, a float or double? Which is more precise?

(c) If I'm trying to establish if two integers are equal, what's the easiest way to compare their values in code? Write the statement that would achieve this?

(d) If I'm trying to establish if two real numbers (double or float) are equal, what's the "correct" way to compare their values in code? Write the statement that would achieve this.

(e) What is Python default machine precision?

Solution:

(a) unsigned int

(b) double, double

(c) a == b

(d) $|a - b| < \epsilon$

(e) Double (64 bits or I will also accept 53 bit if they are talking about the mantissa), 9223372036854775807, (`sys.maxsize`) 1.7976931348623157e+308 (`sys.float_info.max`) [10e+4931 with numpy]

6. Base Conversion (10 pts) By-hand. Show all steps:

(a) Convert 723 from decimal to binary.

(b) Convert 0.1 from decimal to binary using 1 byte (8 bits).

(c) Does adding zeros to the back (ones side) of a base 10 integer change the value? What about a base 2 (binary) integer?

(d) Now consider the value to be a decimal. Does anything change?

Solution:

(a) 1011010011

(b) 0.00011001

(c) Yes, yes

(d) No (the number is unchanged)

7. Error Propagation (20 pts) By-hand. Consider the following two statements,

$$z = 2x - y + \sin(xy^2)$$
$$w = e^z - 2(z - 1)$$

with mean values $\mu_x = 2$ and $\mu_y = 1$, and standard deviations $\sigma_x = 0.03$ and $\sigma_y = 0.01$. Estimate the following parameters using four significant digits:

(a) μ_w

(b) σ_w

Solution:

(a) $\mu_z = 2\mu_x - \mu_y + \sin(\mu_x\,\mu_y^2) = 3.909$
$\mu_w = e^{\mu_z} - 2(\mu_z - 1) = 44.045$

(b) Partial derivatives:

$$\left.\frac{\partial z}{\partial x}\right|_{\mu_x,\mu_y} = 2 + \mu_y^2\,\cos(\mu_x\mu_y^2) = 1.584$$
$$\left.\frac{\partial z}{\partial y}\right|_{\mu_x,\mu_y} = -1 + 2\mu_x\mu_y\,\cos(\mu_x\mu_y^2) = -2.665$$
$$\left.\frac{\partial w}{\partial z}\right|_{\mu_z} = e^{\mu_z} - 2 = 47.85;$$

Then, the standard deviation of z becomes,

$$\sigma_z = \sqrt{\left.\frac{\partial z}{\partial x}\right|^2_{\mu_x,\mu_y}\sigma_x^2 + \left.\frac{\partial z}{\partial y}\right|^2_{\mu_x,\mu_y}\sigma_y^2} = 0.0545$$

and the standard deviation of w is,

$$\sigma_w = \sqrt{\left.\frac{\partial w}{\partial z}\right|^2_{\mu_z}\sigma_z^2} = 2.607$$