Data: 05.03.2025

Laboratorium 1

Autorzy: Kosman Mateusz, Ludwin Bartosz

Temat laboratorium

Celem zajęć było zbadanie wpływu błędów numerycznych na obliczenia, a także opracowanie metod minimalizujących ich skutki, m.in. poprzez stosowanie różnych metod różnicowych, sumowania liczb zmiennoprzecinkowych oraz przekształcania wyrażeń matematycznych w celu uniknięcia kancelacji.

Zadanie 1:

Treść:

Zadanie polegało na obliczeniu pochodnych funkcji trygonometrycznej tangens w punkcie x = 1 za pomocą wzorów na różnice prawostronne oraz centralne:

$$f'(x) \approx \frac{f(x+h) - f(x)}{h} \tag{1}$$

$$f'(x) \approx \frac{f(x+h) - f(x-h)}{2h} \tag{2}$$

a następnie przedstawieniu na wspólnym wykresie wartości bezwzględnej błędu metody, błędu numerycznego oraz błędu obliczeniowego $E(h)$ w zależności od zadanych wartości h dla każdego wzoru z osobna.

Różnica prawostronna:

Błąd metody (truncation error) obliczamy korzystając ze wzoru:

$$E_t = \frac{Mh}{2} \tag{3}$$

gdzie M to przybliżona wartość pochodnej funkcji tangens w punkcie x = 1.

Błąd numeryczny (rounding error) obliczamy używając wzoru:

$$E_r = \frac{2\epsilon_{mach}}{h} \quad (4)$$

gdzie ϵ_{mach} to epsilon maszynowy w języku Python (tzn. minimalna liczba ϵ_{mach} dla której warunek 1 + ϵ_{mach} > 1 jest spełniony). W tym przypadku ta wartość to około $2.220446049250313 \cdot 10^{-16}$.

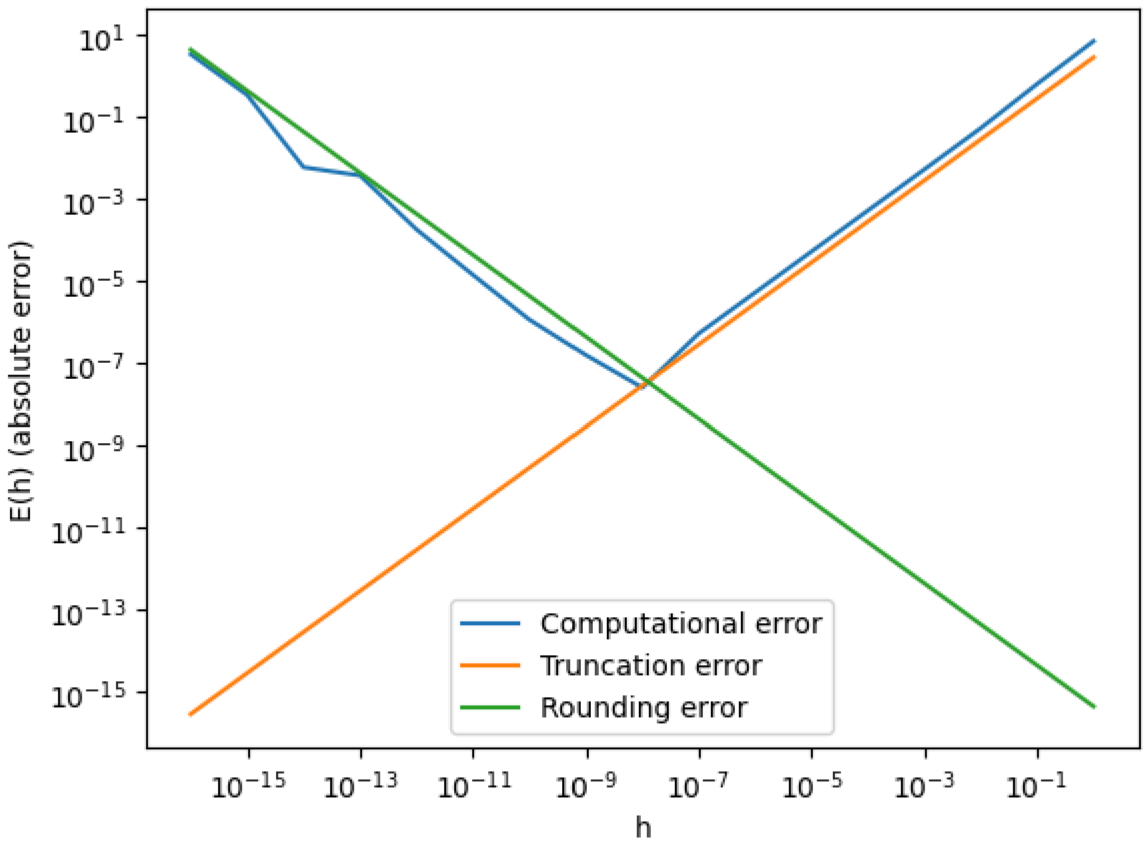
Błąd bezwzględny metody (oznaczany w kodzie abs_error) wyznaczamy ze wzoru:

$$E = |f'(x) - \widehat{f}'(x)| \quad (5)$$

gdzie $f'(x)$ to poprawna wartość pochodnej funkcji tangens w punkcie x = 1 (została ona obliczona za pomocą zależności $tan'(x) = 1 + tan(x)^2$) a poprzez $\widehat{f}'(x)$ oznaczamy metodycznie wyznaczoną wartość pochodnej.

Wykres i wyniki:

Przedstawiając wyniki w skali logarytmicznej na obu osiach otrzymujemy wykres:



Wykres 1 - błędy dla różnicy prawostronnej

Wykres wyraźnie pokazuje że błędy przestrzegają nierówności:

$$E(h) \leq \frac{Mh}{2} + \frac{2\epsilon_{mach}}{h} \tag{6}$$

w której prawa strona jest kombinacją wzorów (2) oraz (3).
Ponadto można zaobserwować minimum funkcji $E(h)$ w punkcie $h = 10^{-8}$.

Zostało ono również obliczone przez program funkcją:

```python
def empirical_h_min(self, abs_difference : np.array) -> float:
        """
        Wyznacza empirycznie krok h dla którego
        wartość bezwzględnej różnicy pomiędzy wartością numeryczną
        a analityczną (podana w abs_difference) jest najmniejsza.
        """
        idx_min = np.argmin(abs_difference)  # Znajdź indeks minimalnego błędu
        return self._h_array[idx_min]          # Zwróć odpowiadający krok h
```

Porównując go z teoretyczną wartością obliczoną ze wzoru:

$$h_{min} = 2\sqrt{\frac{\epsilon_{mach}}{M}} \tag{7}$$

gdzie M to przybliżona wartość drugiej pochodnej funkcji tangens w punkcie x = 1, okazuje się że ich różnica względna to około 19.43%, co jest prawdopodobnie spowodowane niską rozdzielczością punktów na osi h dla wartości bliskich minimum. Dokładna wartość teoretycznego h_{min} została zapisana w tabelce we wnioskach.

Różnica centralna:

W tym wypadku procedura analizy przebiega tak samo. Należy jednak zmodyfikować niektóre wzory na wymienione poniżej:
Błąd metody (truncation error) obliczamy teraz korzystając ze wzoru:

$$E_t = \frac{Mh^2}{6} \tag{8}$$

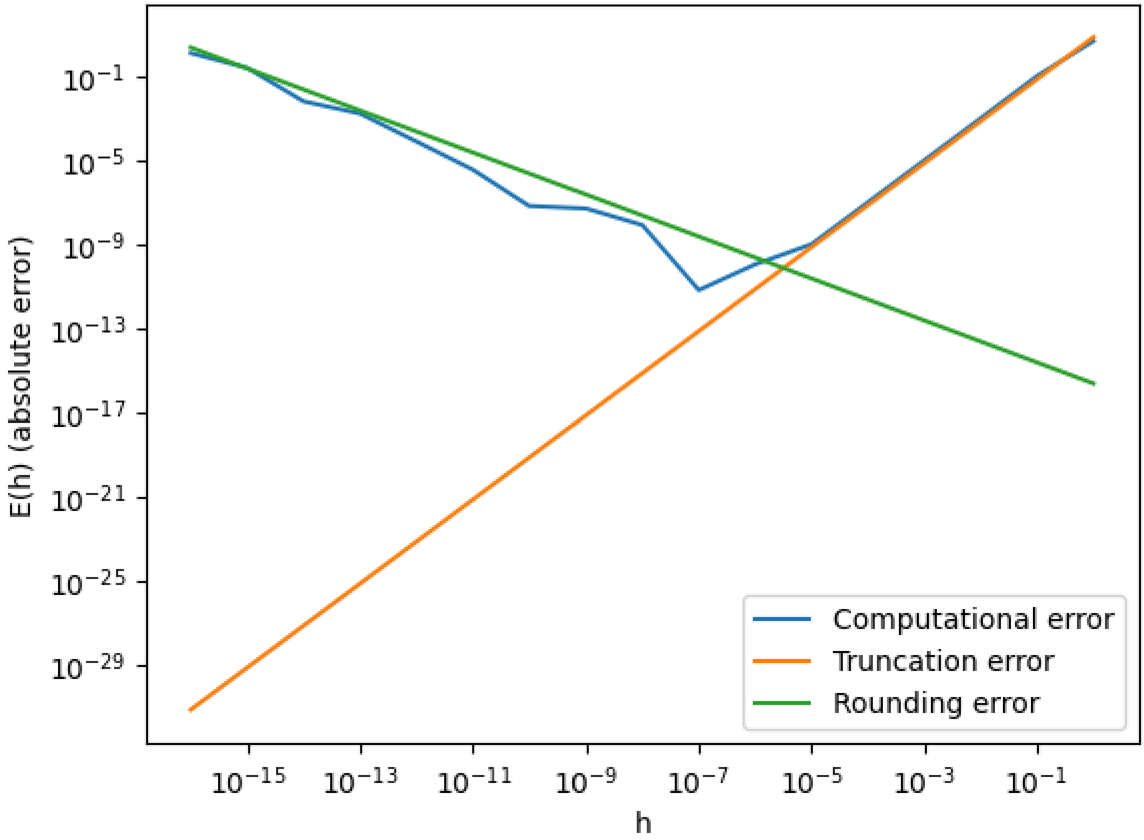
W tym przypadku M oznacza drugą pochodną funkcji tangens w punkcie x= 1.
Błąd numeryczny (rounding error) obliczamy natomiast używając wzoru:

$$E_r = \frac{\epsilon_{mach}}{h} \tag{9}$$

Reszta oznaczeń pozostaje bez zmian.

Wykres i wyniki:

Przedstawiając wyniki w skali logarytmicznej na obu osiach otrzymujemy wykres:



Wykres 2 - błędy dla różnicy centralnej

Wykres wyraźnie pokazuje że błędy przestrzegają następującej nierówności:

$$E(h) \leq \frac{Mh^2}{6} + \frac{\epsilon_{mach}}{h} \tag{10}$$

w której prawa strona jest kombinacją wzorów (8) oraz (9).

Ponadto można zaobserwować minimum funkcji $E(h)$ w punkcie $h = 10^{-7}$.
Zgadza się ono z minimum wyliczonym funkcją `empirical_h_min`.
Porównując go z teoretyczną wartością obliczoną ze wzoru:

$$h_{min} = \sqrt[3]{3\frac{\epsilon_{mach}}{M}} \tag{11}$$

gdzie M oznacza wartość drugiej pochodnej funkcji tangens w punkcie x = 1.
W tym przypadku jednak różnica względna to aż około 96.04%. Dokładna wartość teoretycznego h_{min} została zapisana w tabelce we wnioskach.

Wnioski:

Typ różniczkowania od h_{min}	h_{min} empiryczne	h_{min} teoretyczne	Błąd względny
Różnice prawostronne	10^{-7}	$1.24 \cdot 10^{-8}$	19.43%
Różnice centralne	10^{-7}	$2.52 \cdot 10^{-6}$	96.04%

Porównując wyznaczone wartości błędów obliczeniowych $E(h_{min})$ dla h_{min} wyznaczonych ze wzorów (zapewniają one dokładniejsze wartości) możemy zauważyć że wartość dla różnicy centralnej ($E(h_{min}) \approx 6.22 \cdot 10^{-12}$) jest o 4 rzędy wielkości mniejsza niż dla różnicy prawostronnej ($E(h_{min}) \approx 2.55 \cdot 10^{-8}$). Oznacza to że metoda wykorzystująca wzór na różnice centralne jest dokładniejsza. Dzieje się tak ponieważ reszta we wzorze Taylora dla różnicy prawostronnej jest rzędu $O(h)$, natomiast dla metody wykorzystującej różnice centralne to$O(h^2)$. Powoduje to że dla małych wartości h błąd minimalizuje się szybciej dla różnic centralnych.

Zadanie 2

Treść:

Napisz program obliczający sumę n liczb zmiennoprzecinkowych pojedynczej precyzji, losowo rozłożonych w przedziale [0,1] wg rozkładu jednostajnego. Użyj wyłącznie zmiennych pojedynczej precyzji, chyba, że wskazano inaczej. Sumę oblicz według każdego z poniższych sposobów:

a. Zsumuj liczby według kolejności, w której zostały wygenerowane. Użyj akumulatora podwójnej precyzji do przechowywania akumulowanej sumy.
b. Zsumuj liczby według kolejności, w której zostały wygenerowane. Użyj akumulatora pojedynczej precyzji do przechowywania akumulowanej sumy.
c. Użyj algorytmu Kahana sumowania z kompensacją, sumując liczby w kolejności, w której zostały wygenerowane. Użyj akumulatora pojedynczej precyzji do przechowywania akumulowanej sumy.
d. Zsumuj liczby w porządku rosnącym, od liczb o najmniejszej wartości bezwzględnej do liczb o największej wartości bezwzględnej.
e. Zsumuj liczby w porządku malejącym, od liczb o największej wartości bezwzględnej do liczb o najmniejszej wartości bezwzględnej.

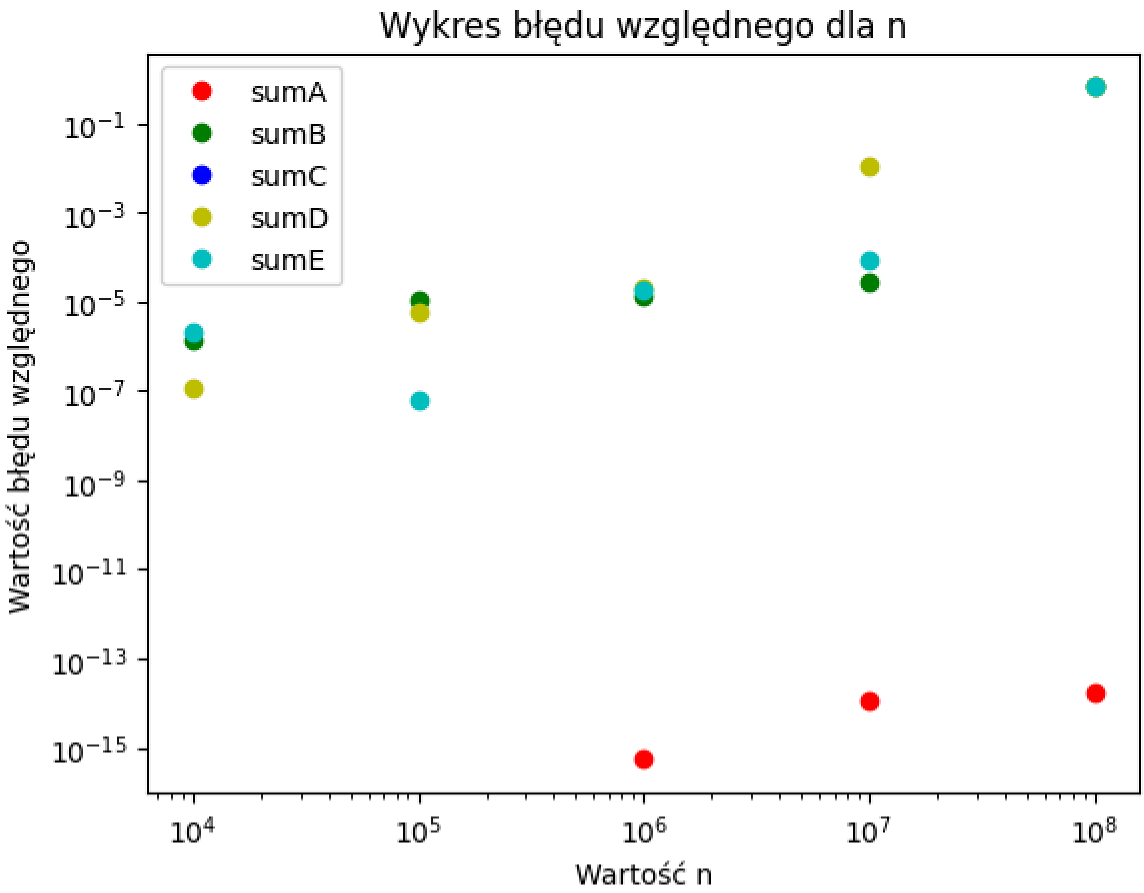
Narysuj wykres błędu względnego w zależności od n = 10^k, k = 4, ..., 8. Jako prawdziwą wartość sumy przyjmij wartość *np.fsum(x)*.

Rozwiązanie:

Definiujemy metody: *sumA, sumB, sumC, sumD, sumE* obliczające sumę zgodnie z definicją podaną w treści zadania. Korzystając z tych funkcji obliczamy błąd względny podanych aproksymacji zakładając, że *math.fsum* zwraca wartość prawdziwą. Wartość błędu względnego obliczamy ze wzoru:

$$\eta = \left|\frac{x-\hat{x}}{x}\right| = \left|1 - \frac{\hat{x}}{x}\right|$$

Wykres przedstawiający błąd względny dla przykładowego wywołania testów:



Wykres 3 - Wykres błędu względnego dla zadanego n

Wnioski:

Sumowanie z wykorzystaniem akumulatora podwójnej precyzji (*np.float64*) oraz wykorzystanie algorytmu Kahana (ze zmiennymi pomocniczymi pojedynczej precyzji - *np.float32*) dają najmniejszy błąd względny względem funkcji *math.fsum()*. Pozostałe metody charakteryzują się mniejszą dokładnością. Dokonując obliczenia ze zmiennymi ograniczonej precyzji warto zatem rozważyć wykorzystanie akumulatora o wyższej precyzji niż dane wejściowe lub wykorzystać algorytm, który charakteryzuje się wyższą precyzją wyniku niż sumowanie naiwne (np. algorytm Kahana).

Zadanie 3

Treść:

Przepisz poniższe wyrażenia, tak aby uniknąć zjawiska kancelacji dla wskazanych argumentów.

a. $\sqrt{x + 1} - 1, x \approx 0$
b. $x^2 - y^2, x \approx y$
c. $1 - \cos x,\ x \approx 0$
d. $\cos^2 x - \sin^2 x,\ x \approx 0$
e. $\ln x - 1,\ x \approx e$
f. $e^x - e^{-x},\ x \approx 0$ (Wskazówka. Użyj rozwinięcia w szereg Taylora).

Rozwiązanie:

a. $\sqrt{x + 1} - 1 = \frac{(x+1)-1}{\sqrt{x+1}+1} = \frac{x}{\sqrt{x+1}+1}$
b. $x^2 - y^2 = (x - y)(x + y)$
c. $1 - \cos x = 1 - (1 - 2\sin^2(\frac{x}{2})) = 2\sin^2(\frac{x}{2})$
d. $\cos^2 x - \sin^2 x = \cos(2x)$
e. $\ln x - 1 = \ln x - \ln e = \ln(\frac{x}{e})$
f. $e^x - e^{-x} \approx (\sum_{k=1}^{5} \frac{x^k}{k!}) - (\sum_{k=1}^{5} \frac{(-x)^k}{k!}) = 2\sum_{k=0}^{2} \frac{x^{2k+1}}{(2k+1)!}$

Wnioski:

Wiele wyrażeń arytmetycznych można przekształcić równoważnie, aby ich obliczanie korzystające ze zmiennych o ograniczonej precyzji minimalnie traciło precyzję rozwiązania. Można w tym celu wykorzystać m.in. wzory skróconego mnożenia (np. a., b.), tożsamości trygonometryczne (np. c., d.), własności logarytmów (np. e.) oraz rozwinięcie wyrażenia w szereg Taylora (np. f.).

Zadanie 4:

Treść:

Celem zadania było stwierdzenie czy możemy być pewni że kolektor słoneczny $S1$ ma większą sprawność niż kolektor $S2$. Sprawność kolektora dana jest wzorem:

$$\eta = K \frac{QT_d}{I} \tag{12}$$

gdzie K jest stałą znaną z dużą dokładnością, Q – objętość przepływu, T_d - różnica temperatur, I – natężenia promieniowania.

Wyliczona wartość sprawności ze wzoru (12) dla ogniwa $S1$ wynosi 0.76, natomiast dla $S2$ jest to 0.70. Wielkości Q, T_d oraz I zmierzono z następującymi błędami:

Kolektor / Wielkość	$S1$	$S2$
Q	1.5%	0.5%
T_d	1.0%	1.0%
I	3.6%	2.0%

Rozwiązanie

Aby odpowiedzieć na stwierdzenie z treści należy oszacować maksymalny możliwy błąd (tzw. worst-case), a następnie sprawdzić czy istnieje możliwość aby $S1$ było mniejsze niż $S2$ w granicach niepewności. W tym celu używamy wzoru:

$$\frac{\Delta\eta}{\eta} = \frac{\Delta Q}{Q} + \frac{\Delta T_d}{T_d} + \frac{\Delta I}{I} \tag{13}$$

Dla ogniwa $S1$ będzie to:

$$\frac{\Delta\eta_1}{\eta_1} = 1.5\% + 1.0\% + 3.6\% = 6.1\%$$

Natomiast dla $S2$:

$$\frac{\Delta\eta_2}{\eta_2} = 0.5 + 1.0\% + 2.0\% = 3.5\%$$

Przedziały ufności dla poszczególnych kolektorów przedstawiają się zatem następująco:

$$S1:\ [0.76\ -\ 0.061,\ 0.76\ +\ 0.061]\ =\ [0.699,\ 0.821]$$

$$S2:\ [0.70\ -\ 0.035,\ 0.70\ +\ 0.035]\ =\ [0.665,\ 0.735]$$

Przedziały dla $S1$ oraz $S2$ posiadają część wspólną ($[0.699, 0.735]$), więc nie możemy być pewni że $S1$ ma większą wartość niż $S2$.

Wnioski:

Analiza powyższego przykładu pokazuje istotność niepewności pomiarowych w rzetelnej ocenie danych. Niedoszacowanie błędów lub ich całkowite pominięcie mogłoby prowadzić do błędnych interpretacji uzyskanych wyników. Uwzględnienie przedziałów niepewności pozwala nie tylko na określenie zakresu wiarygodności mierzonych wartości, a co za tym idzie – zachowaniu obiektywności oceny.

Wnioski ogólne:

Stosowanie odpowiednich metod (np. różnica centralna, algorytm Kahana) znacząco poprawia dokładność obliczeń, a właściwe uwzględnienie niepewności pomiarowych jest kluczowe dla rzetelnej oceny wyników. Laboratorium pokazało że poprzez świadomy dobór technik można zminimalizować utratę precyzji i lepiej interpretować otrzymane rezultaty.

Źródła:

- https://pf.agh.edu.pl/pomoce-dydaktyczne
- Prezentacje z MS Teams (zespół MOwNiT 2025)