Laboratorium 1

Arytmetyka Komputerowa

Dominik Marek

5 marca 2024

AGH

Akademia Górniczo-Hutnicza
im. Stanisława Staszica
w Krakowie

1. Treść zadań

1. Znaleźć "*maszynowe epsilon*", czyli najmniejszą liczbę *a*, taką że *a*+1>1
2. Rozważamy problem ewaluacji funkcji sin(x), m.in. propagację błędu danych wejściowych, tj. błąd wartości funkcji ze względu na zakłócenie h w argumencie x:
    a) Ocenić błąd bezwzględny przy ewaluacji sin(x)
    b) Ocenić błąd względny przy ewaluacji sin(x)
    c) Ocenić uwarunkowanie dla tego problemu
    d) Dla jakich wartości argumentu x problem jest bardzo czuły ?

3. Funkcja sinus zadana jest nieskończonym ciągiem

$$sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots$$

a) Jakie są błędy progresywny i wsteczny jeśli przybliżamy funkcję sinus biorąc tylko pierwszy człon rozwinięcia, tj. sin(x) ≈ x, dla x = 0.1, 0.5 i 1.0 ?

b) Jakie są błędy progresywny i wsteczny jeśli przybliżamy funkcję sinus biorąc pierwsze dwa człony rozwinięcia, tj. $sin(x) \approx x - \frac{x^3}{6}$, dla x = 0.1, 0.5 i 1.0 ?

4. Zakładamy że mamy znormalizowany system zmiennoprzecinkowy z
β = 10, p = 3, L = -98 .

a) Jaka jest wartość poziomu UFL (underflow) dla takiego systemu ?
b) Jeśli x = $6{,}87 \cdot 10^{-97}$ i y = = $6{,}81 \cdot 10^{-97}$ jaki jest wynik operacji x – y ?

2.Rozwiązania

1.

Maszynowe epsilon to największa liczba nieujemna, której dodanie do jedności daje wynik równy 1. Zapisując inaczej $1 + \epsilon = 1$. Maszynowe epsilon musi mieć identyczny jak cyfra 1 jak również najmniejszą mantysę, czyli 1.Zatem maszynowe epsilon jest dane następującym wzorem:

$$\varepsilon = B^{1-p}$$

Gdzie:

B - podstawą systemu liczbowego dla danego epsilon

p - zadana precyzja

Przykładowy kod z języka Python, wyliczający epsilon dla zadanego systemu i precyzji:

```python
def find_machine_epsilon(basis: int, precision: int) -> float:
    return basis**(1-precision)
```

Wynik działania powyższej funkcji dla systemy dwójkowego z precyzją równą 11:

```
Machine epsilon for binary system w precision equal 11  = 0.0009765625
```

2.

a) Błąd bezwzględny obliczamy przy pomocy wzoru :

$$\Delta \sin x = |\sin x\,(1+\varepsilon) - \sin x|$$

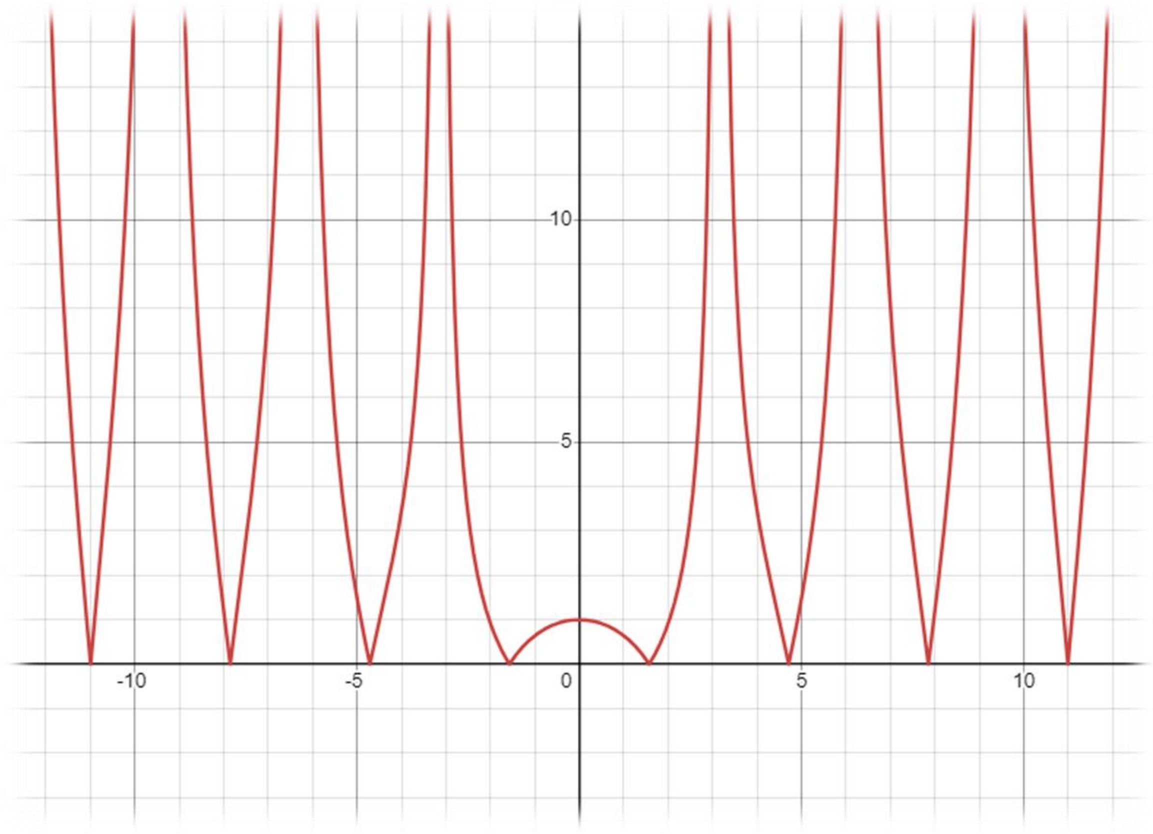
b) Błąd względny obliczamy ze wzoru:

$$\frac{\Delta f(x)}{f(x)} = \frac{\Delta \sin x}{\sin x} = \frac{|\sin(1+\varepsilon)x - \sin x|}{\sin x}$$

c) Uwarunkowanie dane jest następującym wzorem:

$$cond\big(f(x)\big) = \lim_{\hat{x}\to x} \frac{\left|\frac{f(x) - f(\hat{x})}{f(x)}\right|}{\left|\frac{x-\hat{x}}{x}\right|} = \left|\frac{x \cdot f'(x)}{f(x)}\right|$$

$$cond(\sin x) = \lim_{\hat{x}\to x} \frac{\left|\frac{\sin x - \sin(1+\varepsilon)x}{\sin x}\right|}{\left|\frac{x-(1+\varepsilon)x}{x}\right|} = \left|\frac{x \cdot \sin^7(x)}{\sin x}\right| = \left|\frac{x\cos x}{\sin x}\right| = |x\cot x|$$



Rysunek 1: Wykres funkcji $f(x) = |x \cot x|$

d) Jak można zauważyć z wyżej wyprowadzonego wzoru na uwarunkowanie funkcji $sin\, x$ oraz powyższego wykresu, problem jest bardzo czuły w punktach zadanych równaniem $x = k\pi, k \in \mathbb{Z} / \{0\}$, dla których funkcja uwarunkowania dąży do nieskończoności. Natomiast najlepiej uwarunkowany jest dla punktów zerowania się funkcji $cos\, x$, czyli $x = \frac{k\pi}{2}$, $k \in \mathbb{Z}$.

Wnioski: Funkcja sinus jest najlepiej uwarunkowana , w punktach gdzie osiąga swoje minimum , czyli wartość 0, natomiast w punktach, gdzie przyjmuje swoje maksimum równe 1, jest najgorzej uwarunkowana. Własność ta wynika z faktu, iż funkcja sinus ma osiąga swoje ekstrema w punktach, będących miejscami zerowymi funkcji cosinus i analogicznie zachowuje się cosinus.

3.

Błąd progresywny to wartość bezwzględna z różnicy wartości rzeczywistej i uzyskanej (przybliżonej). Błąd wsteczny to z kolei wartość bezwzględna z różnicy argumentu wstawionego do funkcji i argumentu, dla którego przybliżona wartość funkcji jest wartością rzeczywistą. Badamy funkcję postaci :

$$y = \sin x$$

a) Przybliżamy funkcję jako $sin(x) \approx x$

$\hat{y} = x$

$\hat{x} = arc\ sin\, \hat{y}$

Błąd progresywny obliczamy ze wzoru:

$$|y - \hat{y}| = |sin\, x - x|$$

Błąd wsteczny dany jest następującym wyrażeniem:

$$|\hat{x} - x| = |arcsin\, \hat{y} - x| = |arcsin\, x - x|$$

Rachunki:

- *Dla* $x = 0.1$
  $\hat{y} = x = 0{,}1$
  $\hat{x} = arcsin\, \hat{y}$ *= arcsin(x) ≈ 0.1001674212*
  Błąd progresywny: |sin (0.1) − 0.1| ≈ 0.0001665834
  *Błąd wsteczny: |arcsin(*0.1*) − 0.1| ≈ 0.0001674212*

- *Dla* $x = 0.5$
  $\hat{y} = x = 0{,}5$
  $\hat{x} = arcsin\, \hat{y}$ *= arcsin(x) ≈ 0.5235987756*
  Błąd progresywny: |sin (0.5) − 0.5| ≈ 0.0205744614
  *Błąd wsteczny: |arcsin(*0.5*) − 0.5| ≈* 0.0235987756

- *Dla* $x = 1.0$
  $\hat{y} = x = 1.0$
  $\hat{x} = arcsin\, \hat{y}$ *= arcsin(x) ≈ 1.5707963268*
  Błąd progresywny: |sin (1.0) – 1,0| ≈ 0.1585290152
  *Błąd wsteczny: |arcsin(*1.0*) – 1.0| ≈ 0. .5707963268*

b) Przybliżamy funkcję jako $sin(x) \approx x - \frac{x^3}{6}$

$$\hat{y} = x - \frac{x^3}{6}$$

$$\hat{x} = arcsin(x - \frac{x^3}{6})$$

Błąd progresywny obliczamy ze wzoru:

$$|y - \hat{y}| = \left|sin\, x - (x - \frac{x^3}{6})\right| = \left|sin\, x + \frac{x^3}{6} - x\right|$$

Błąd wsteczny dany jest następującym wyrażeniem:

$$|\hat{x} - x| = |\arcsin \hat{y} - x| = \left|arcsin(x - \frac{x^3}{6}) - x\right|$$

Rachunki:

- *Dla* $x = 0.1$
  $\hat{y} = x - \frac{x^3}{6} = 0.1 - \frac{(0.1)^3}{6} \approx 0.0998333332$
  $\hat{x} = arcsin(x - \frac{x^3}{6}) \approx 0.0999999163$

  Błąd progresywny: $|sin\ 0.1 - 0.\ 0998333332| \approx 0.000000083$

  Błąd wsteczny: $|arcsin(0.0998333332) - 0.1| \approx 0.000000084$

- *Dla* $x = 0.5$
  $\hat{y} = x - \frac{x^3}{6} = 0.5 - \frac{(0.5)^3}{6} \approx 0.4791666667$
  $\hat{x} = arcsin(x - \frac{x^3}{6}) \approx 0.4997050408$

  Błąd progresywny: $|sin\ 0.5 - 0.4791666667| \approx 0.0002588719$

  Błąd wsteczny: $|arcsin(0.479167000) - 0.5| \approx 0.0002949592$

- *Dla* $x = 1.0$
  $\hat{y} = x - \frac{x^3}{6} = 1.0 - \frac{(1{,}0)^3}{6} \approx 0.8333333332$
  $\hat{x} = arcsin(x - \frac{x^3}{6}) \approx 0.9851107832$

  Błąd progresywny: $|sin\ 1.0 - 0.833333332| \approx 0.0081376515$

  Błąd wsteczny: $|arcsin(0.833333332) - 0.1| \approx 0.0148892167$

Wnioski: Rozszerzając funkcję sinus w szereg Taylora przy większej liczbie uwzględnianych wyrazów otrzymujemy wyniki bliższe rzeczywistej wartości. Możemy na tej podstawie wnioskować, że dokładność wyników rośnie wraz z liczbą wykorzystanych wyrazów szeregu.

4.

a) Poziom UFL to najmniejsza liczba dodatnia, jaka może zostać zapisana w danym systemie. Możemy zauważyć, że mantysa tej liczby musi być równa 1, zaś jej wykładnik powinien być możliwie jak najmniejszy. Zatem w naszym przypadku otrzymujemy:

$$UFL = \beta^L = 10^{-98}$$

b) Dla $x = 6{,}87 \cdot 10^{-97}$ i $y = 6{,}81 \cdot 10^{-97}$

otrzymujemy różnicę x-y $= 0{,}06 \cdot 10^{-97} = 6{,}00 \cdot 10^{-99} < UFL$

Możemy zauważyć, że otrzymana różnica jest mniejsza niż UFL, a zatem wynik tej operacji będzie wynosił 0.

Wnioski: Wiedząc, że UF L jest miarą dokładności systemu zmiennoprzecinkowego to chcą uzyskać jak najbardziej dokładny system operujący na małych liczbach należy zminimalizować parametr L.

3.Bibliografia:

- Katarzyna Rycerz: Wykład z przedmiotu Metody Obliczeniowe w Nauce i Technice
- Michael T. Heath: Scientific Computing: An Introductory Survey
- https://en.wikipedia.org/wiki/IEEE_754-1985
- https://en.wikipedia.org/wiki/Machine_epsilon