Laboratorium 1

Arytmetyka komputerowa

Jan Rajczyk

12 marca 2021

1 Treści 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

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + ...$$

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

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

4. Zakładamy że mamy znormalizowany system zmiennoprzecinkowy z $\beta = 10$, $p = 3$, $L = -98$:

   (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

Obliczenia zostały wykonane z użyciem programu WolframAlpha, zaś wykresy z użyciem programu Desmos.

1. *Maszynowe epsilon* musi mieć taki sam wykładnik jak liczba 1 oraz jak najmniejszą mantysę, a więc równą 1. Stąd *maszynowe epsilon* jest równe:

$$\epsilon = B^{1-p},$$

gdzie p jest precyzją, zaś B podstawą systemu liczbowego.

2.
a) Błąd bezwględny

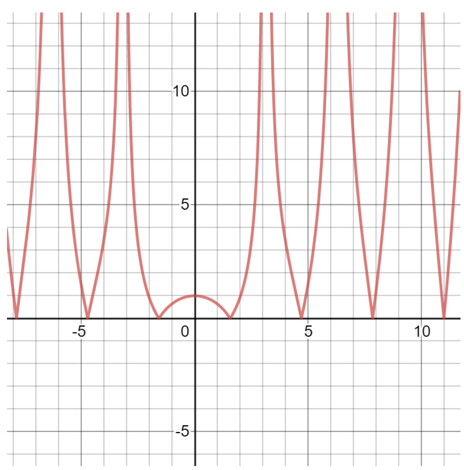
$$\Delta \sin x = |\sin x(1+\epsilon) - \sin x|$$

b) Błąd względny

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

c) Uwarunkowanie

$$cond(f(x)) = cond(\sin x) = |\frac{xf'(x)}{f(x)}| = |\frac{x\cos x}{\sin x}| = |x \cot x|$$



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

d) Problem jest czuły w miejscach, gdzie funkcja $\cot x$ zmierza do nieskończoności, czyli dla $x = k\pi$, gdzie $k \in Z$, poza $k = 0$.
Najlepiej uwarunkowany natomiast będzie problem w miejscach, gdzie funkcja $\cos x$ zeruje się, a więc dla $x = k\pi + \pi/2$, gdzie $k \in Z$.

Wnioski: Możemy zauważyć, że funkcja sinus jest najgorzej uwarunkowana w swoich miejscach zerowych, zaś najlepiej w miejsach, gdzie przyjmuje wartość równą 1.
Jest to oczywiście związane z tym, że odpowiednio dla miejsc zerowych funkcji sinus cosinus przyjmuje wartości ekstremalne, zaś w eskremach funkcji sinus cosinus przyjmuje wartości równe 0.

3. Błąd progresywny to wartość bezwględna z różnicy wartości rzeczywistej i uzyskanej (przybliżonej). Błąd wsteczny to z kolei wartość bezwlędna z różnicy argumentu wstawionego do funkcji i argumentu, dla którego przybliżona wartość funkcji jest wartośćią rzeczywistą.
Będziemy rozpatrywać funkcję postaci:

$$y = \sin x$$

a) $\hat{y} = x$, $\hat{x} = \arcsin \hat{y}$
Błąd progresywny: $|y - \hat{y}| = |\sin x - x|$
Błąd wsteczny: $|\hat{x} - x| = |\arcsin \hat{y} - x| = |\arcsin x - x|$
Obliczenia:

- $x = 0.1$:
  $\hat{y} = x = 0.1$
  $\hat{x} = \arcsin \hat{y} \approx 0.100167421$
  Błąd progresywny: $|\sin 0.1 - 0.1| \approx 0.0982546163$
  Błąd wsteczny: $|\hat{x} - 0.1| \approx 0.000167421$
- $x = 0.5$:
  $\hat{y} = x = 0.5$
  $\hat{x} = \arcsin \hat{y} \approx 0.523598776$
  Błąd progresywny: $|\sin 0.5 - 0.5| \approx 0.020574461$
  Błąd wsteczny: $|\hat{x} - 0.5| \approx 0.023598776$
- $x = 1.0$:
  $\hat{y} = x = 1.0$
  $\hat{x} = \arcsin \hat{y} \approx 1.570796327$
  Błąd progresywny: $|\sin 1.0 - 1.0| \approx 0,158529015$
  Błąd wsteczny: $|\hat{x} - 1.0| \approx 0.570796327$

b) $\hat{y} = x - \frac{x^3}{6}$, $\hat{x} = \arcsin (x - \frac{x^3}{6})$
Błąd progresywny: $|y - \hat{y}| = |\sin x - (x - \frac{x^3}{6})| = |\sin x + \frac{x^3}{6} - x|$
Błąd wsteczny: $|\hat{x} - x| = |\arcsin \hat{y} - x| = |\arcsin (x - \frac{x^3}{6}) - x|$
Obliczenia:

- $x = 0.1$:
  $\hat{y} = 0.1 - \frac{(0.1)^3}{6} \approx 0.099833333$
  $\hat{x} = \arcsin \hat{y} \approx 0.099999917$
  Błąd progresywny: $|\hat{y} - 0.1| \approx 0.000000083$
  Błąd wsteczny: $|\hat{x} - 0.1| \approx 0.000000084$
- $x = 0.5$:
  $\hat{y} = 0.5 - \frac{(0.5)^3}{6} \approx 0.479167000$
  $\hat{x} = \arcsin \hat{y} \approx 0.499705041$
  Błąd progresywny: $|\hat{y} - \frac{(0.5)^3}{6}| \approx 0.000258872$
  Błąd wsteczny: $|\hat{x} - 0.5| \approx 0.000294959$
- $x = 1.0$:
  $\hat{y} = 1.0 - \frac{(1.0)^3}{6} \approx 0.833333332$
  $\hat{x} = \arcsin \hat{y} \approx 0.985110783$
  Błąd progresywny: $|\hat{y} - \frac{(1.0)^3}{6}| \approx 0.008137652$
  Błąd wsteczny: $|\hat{x} - 1.0| \approx 0.014889216$

Wnioski: Zauważmy, że przy użyciu rozszerzenia funkcji sinus z szeregu Taylora z większą ilością wyrazów otrzymujemy nieco bardziej dokładne ładne wyniki. Może to sugerować, że im więcej wyrazów z tego szeregu użyjemy to tym lepszą dokładność otrzymamy.

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. Otrzymujemy więc:

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

b) Zauważmy, że $x - y = 0.06 \cdot 10^{-97} = 6 \cdot 10^{-99} < UFL$, stąd w tym systemie wynik operacji $x - y$ będzie wynosił 0.
Wnioski: Z racji tego, że UFL jest miarą dokładności systemu zmiennoprzecinkowego to system operujący na małych liczbach chcąc być najbardziej dokładnym powinien mieć jak najmniejszy parametr L.

3 Bibliografia

- Katarzyna Rycerz: *Wykład z przedmiotu Metody Obliczeniowe w Nauce i Technice*
- `https://en.wikipedia.org/wiki/IEEE_754-1985`
- `https://en.wikipedia.org/wiki/Machine_epsilon`