Laboratorium 1

Arytmetyka komputerowa

Krzysztof Solecki

4 Marca 2023

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

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

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

Jakie są błędy progresywny i wsteczny jeśli przybliżamy funkcję sinus biorąc pierwsze dwa człony rozwinięcią, 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 $\beta = 10$, p = 3, L = -98

   (a) Jaka jest wartość poziomu UFL (underflow) dla takiego systemu ?

   (b) Jeśli $x = 6.87 \times 10^{(-97)}$ i $y = 6.81 \times 10^{(-97)}$, jaki jest wynik operacji x − y ?

2 Rozwiązania

Obliczenia zostały wykonane z użyciem programu WolframAlpha, a wykresy za pomocą programu Desmos.

1. W reprezentacji IEEE 754 kiedy liczby są dodawane najpierw są sprowadzane do wspólnego wykładnika a następnie dodawane są mantysy. Tak więc najmniejsza liczba którą można dodać do jedynki musi mieć ten sam wykładnik co 1, ale jak najmniejszą mantysę. W takim wypadku:

$$\epsilon = \beta^{(1-p)}$$

, gdzie β jest podstawą systemu liczbowego, a p jest precyzją.

2.

a) Ocenić błąd bezwzględny przy ewaluacji sin(x):

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

b) Ocenić błąd bezwzględny przy ewaluacji sin(x):

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

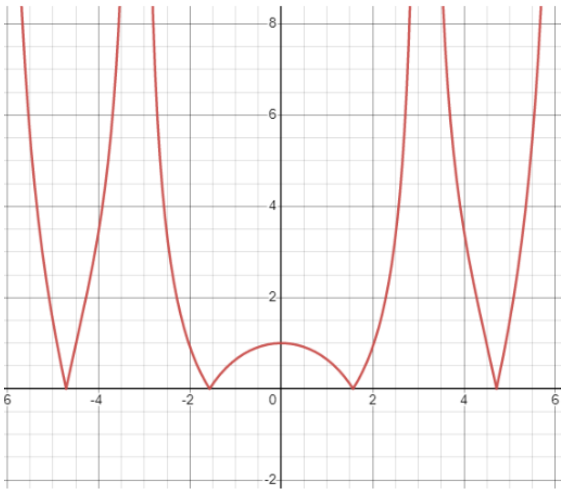
c) Ocenić uwarunkowanie tego problemu:

Dla funkcji jednej zmiennej wzór na uwarunkowanie jest następujący:

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

W przypadku funkcji f(x) = sin(x) otrzymujemy:

$$cond(sin(x)) = |\frac{xcos(x)}{sin(x)}| = |xctg(x)|$$



Rys. 1: Wykres funkcji f(x) = |xctg(x)|

d) Dla jakich wartości parametru x problem jest bardzo czuły:

Problem jest bardzo czuły wtedy, gdy argument x = $k\pi$,gdzie $k \in Z$. Wówczas wartości funkcji ctg(x) zmierzają do nieskończoności (za wyjątkiem k = 0).

Wniosek: Zgodnie z przewidywaniami funkcja sinus jest najgorzej uwarunkowana w otoczeniu swoich miejsc zerowych, gdyż nie dość, że przyjmuje tam bardzo małe wartości to ma największą pochodną. Za to bardzo dobrze jest ona uwarunkowana w miejscach gdzie osiąga ekstrema (największa wartość i pochodna zerująca się).

3. Błąd progresywny to moduł różnicy między wartością uzyskaną a rzeczywistą, natomiast błąd wsteczny to moduł różnicy miedzy wartością podstawioną do funkcji a argumentem dla którego rzeczywista wartość tej funkcji jest równa wartości uzyskanej przez ów przybliżenie funkcji.

Rozpatrujemy funkcję y=sin(x):

x	y = sinx	$\hat{y}$ = sinx ≈ x	błąd progresywny $\|\hat{y} = y\|$	$\hat{x} = arcsin(\hat{y})$	błąd wsteczny $\|\hat{x} - x\|$
0,1	0.099833417	0,1	0.000166583	0,100167421	0,000167421
0,5	0.479425539	0,5	0.020574461	0,523598776	0,023598776
1	0.841470985	1	0.158529015	1,570796327	0,570796327

Tabela 1: Wartości błędów dla przybliżenia sinx = x

x	y = sinx	$\hat{y}$ = sinx ≈ $x - \frac{x^3}{6}$	błąd progresywny $\|\hat{y} = y\|$	$\hat{x} = arcsin(\hat{y})$	błąd wsteczny $\|\hat{x} - x\|$
0,1	0,099833417	0,100166667	0,00033325	0,100334929	0,000334929
0,5	0,479425539	0,520833333	0,041407795	0,547826851	0,047826851
1	0,841470985	0,833333333	0,008137651	0,985110783	0,014889217

Tabela 2: Wartości błędów dla przybliżenia $sin(x) = x - \frac{x^3}{6}$

Wnioski: Zastosowanie większej ilości wyrazów szeregu Taylora powoduje wyraźne zmniejszenie błędów zarówno wstecznych jak i progresywnych. W obu przypadkach błędy rosną co pokrywa się z wynikami z poprzedniego zadania dotyczącymi uwarunkowania funkcji sin(x). Można sądzić zatem, że im więcej wyrazów szeregu użyjemy tym lepszą dokładność wyniku uzyskamy.

4.

a) Poziomem UFL nazywamy najmniejszą liczbę dodatnią jaką można uzyskać w danym systemie. Aby uzyskać taką wartość mantysa musi być równa 1, a wykładnik musi być możliwie jak najmniejszy.

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

b) Można zauważyć, że wynik operacji $x - y = 6 \times 10^{-99} < UFL$, zatem wynik działania x-y w tym systemie liczbowym będzie wynosił 0.

Wnioski: UFL jest miarą dokładności systemu zmiennoprzecinkowego. System służący obliczeniom na bardzo małych liczbach powinien mieć jak najmniejszy parametr L (długość wykładnika charakteryzuje zakres możliwych do uzyskania liczb a mantysa ich dokładność w reprezentacji).

3 Bibliografia

1. https://pl.wikipedia.org/wiki/IEEE$_7$54
2. https://en.wikipedia.org/wiki/Machine$_e psilon$
3. Katarzyna Rycerz: Materiały wykładowe z przedmiotu Metody Obliczeniowe w Nauce i Technice