Uniwersytet Zielonogórski

Laboratorium Metod Numerycznych

Błędy związane z obliczeniami numerycznymi, konwersje liczb

Zakres materiału:

- Algorytmy konwersji liczb pomiędzy systemami zapisu o różnych postawach
- Maszynowa reprezentacja liczb rzeczywistych i całkowitych – format zmiennoprzecinkowy (pojęcie mantysy, cechy, zakresu oraz precyzji) i stałoprzecinkowy
- Algorytm konwersji liczby do postaci maszynowej znormalizowanej stało- i zmiennoprzecinkowej
- Dokładność (precyzja) maszyny cyfrowej (ε)
- Definicja błędu bezwzględnego i względnego
- Pojęcie błędu zaokrąglenia (*roundoff*) oraz obcięcia (*truncation*)
- Pojęcie niedomiaru i nadmiaru zmiennoprzecinkowego (*underflow, overflow*)
- Pojęcie rozwinięcia funkcji w szereg Taylora

Zadania:
Uwaga - w celu rozwiązania zadań 1–8 należy napisać funkcje Matlaba implementujące algorytmy konwersji, pomocne mogą być standardowe funkcje konwertujące Matlaba: *dec2bin, dec2base, dec2hex, bin2dec, hex2dec, base2dec.*

1. Zapisać w systemie dwójkowym, ósemkowym (oktalnym) i szesnastkowym (heksadecymalnym) liczby systemu dziesiętnego:
   a) 24, b) 232, c) 1025, d) $4^6 - 1$, *e*)125,625, d) $0,325$.
2. Zmienić zapis z dziesiętnego na ósemkowy i szestnastkowy:
   a) 16, b) 157, c) 2044.
3. Dokonać następujących konwersji:
   a) $(101101110110)_2 \longrightarrow (?)_8$, b) $(110101010110)_2 \longrightarrow (?)_{16}$, c) $(2716)_8 \longrightarrow (?)_{16}$,
   d) $(\mathrm{F2A})_{16} \longrightarrow (?)_8$.
4. Zapisać dziesiętnie:
   a) $(100111)_2$, b) $(111001001101)_2$, c) $(77)_8$, d) $(263)_8$, e) $(\mathrm{7F})_{16}$, f) $(\mathrm{F8FE})_{16}$, g) $(\mathrm{1A6, E2})_{16}$,
   h) $(77, 44)_8$, i) $(111101, 101)_2$.
5. Przy założeniu, że zapisujemy liczbę w pamięci operacyjnej binarnie w zapisie stałopozycyjnym ze znakiem, podać zakres liczb, które możemy przedstawić na 10 bitach. Jaki będzie zakres dla zapisu zmiennopozycyjnego przy założeniu, że mantysa ma 5 bitów?
6. Przedstawić liczbę $-245, 25$ w zapisie stałopozycyjnym, a następnie zmiennnopozycyjnym przy założeniu, że bazą systemu jest liczba 2. Ile wynosi minimalna liczba bitów potrzebna do przechowania tej liczby w pamięci?
7. Jaką liczbę dziesiętną reprezentują liczby maszynowe (t=4, w=2, b=2, gdzie b - baza, t- liczba cyfr mantysy, w - liczba cyfr cechy bez znaku):
   a) (1)1101(0)10, b) (0)1001(0)00, c) (0)1111(0)11, d) (0)1000(1)11, e) (1)1001(1)01,
8. W wyniku wystąpienia zjawiska zaokrąglania liczba $0, 2$ została zapisana w pamięci maszynowej jako $0, 1875$. Wyznaczyć popełniony błąd bezwzględny i względny. Czy można oszacować na podstawie policzonych błędów dokładność reprezentacji zmiennoprzecinkowej użytej maszyny cyfrowej (ilość bitów mantysy)?

9. **Spontaniczna generacja cyfr nieznaczących**
Należy wykonać ciąg poleceń Matlaba:

```
>> format long e
>> 2.6 + 0.2
>> ans + 0.2
>> ans + 0.2
>> 2.6 + 0.6
```

Wyjaśnić przyczynę pojawienia się błędnej cyfry na najmniej znaczącym miejscu.

10. **Arytmetyka zmiennoprzecinkowa**
Należy porównać efekty wykonania dwóch ciągów poleceń Matlaba:

```
>> format long e
>> u=29/13
>> v=29-13*u
```

oraz

```
>> x=29/1300
>> y=29-1300*x
```

Z czego wynika różnica w wyniku dla pierwszego i drugiego przykładu?

Następnie wprowadzić i zinterpretować wyniki wykonania następujących poleceń Matlaba (jaka jest przyczyna błędu?):

```
>> maks=realmax
>> maks=2*maks
>> minim=realmin
>>minim=minim/1e16
```

Czemu można wykonać bez wystąpienia błędu niedomiaru i jak zinterpretować wynik wykonania operacji:

```
>>realmin/1e14
>>realmin/1e15
```

11. **Precyzja maszynowa**
Wartość popełnianego błędu zaokrąglenia jest limitowana dostępną dla danej maszyny wartością precyzji ε, definowanej jako taka liczba ε, dla której $\underset{\delta<\varepsilon}{\forall}\ 1+\delta=1$. W programie Matlab wartość ta jest dostępna w zmiennej predefiniowanej

```
>>eps
```

Wartość tę można również znaleźć iteracyjnie w sposób przybliżony, korzystając z definicji - startujemy od pewnej wartości poszukiwanej zmiennej, np. $epsilon = 1$, a następnie dzielimy ją na pół tak długo, aż po dodaniu do jedności wynik nie zmieni się (tzn $1 + epsilon = 1$). Należy napisać skrypt Matlaba wyznaczający precyzję maszynową wykonywanych obliczeń.

12. **Błąd obcięcia, błąd względny i bezwzględny**
Rozwinięcie funkcji $\sin(x)$ w szereg Taylora ma postać:

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

Dla małych wartości zmiennej x ($x < 1$) można aproksymować wartość funkcji $\sin(x) \approx x$ (obcięcie rozwinięcia w szereg do pierwszego wyrazu). Należy napisać skrypt Matlaba rysujący wykres błędu względnego i bezwzględnego popełnianego podczas takiej aproksymacji w zakresie $x \in [-0.3\,..\,0.3]$ przy założeniu, że bierzemy pod uwagę odpowiednio pierwsze 1, 2 i 3 człony rozwinięcia funkcji w szereg. Wskazówka – błąd bezwzględny definujemy jako $\Delta = |\hat{\alpha} - \alpha|$, gdzie: $\hat{\alpha}$ – wartość obliczona, α – wartość rzeczywista; bład względny definujemy jako $\delta = \frac{\Delta}{\alpha}$. Dla obcięcia rozwinięcia funkcji $\sin(x)$ do pierwszego wyrazu:

$$\Delta = x - \sin(x) = \frac{x^3}{3!} - \frac{x^5}{5!} + \dots \qquad \delta = \frac{x - \sin(x)}{\sin(x)} = \frac{x}{\sin(x)} - 1$$