Metody numeryczne

dr hab. Piotr Fronczak

Zakład Fizyki Układów Złożonych

- www.if.pw.edu.pl/~agatka/numeryczne.html
- fronczak@if.....
- pok. 101 GF

Regulamin

- Obecność na wykładach nie jest obowiązkowa.
- Zaliczenie wykładów ma formę 2 pisemnych testów, obejmujących zakres materiału, omawiany na wykładach. Testy zapowiadane są z co najmniej dwutygodniowym wyprzedzeniem i organizowane na 7. i 14. zajęciach, chyba że wykładowca w porozumieniu ze studentami postanowi inaczej. Czas trwania każdego testu wynosi 1 godzinę lekcyjną (45 min.)
- Na ostatnich zajęciach organizowany jest dodatkowy test dla osób, które nie zaliczyły jednego lub dwóch testów w terminach podstawowych, lub do nich nie przystąpiły z powodów usprawiedliwionych.
- W trakcie testów można korzystać z własnych notatek z wykładu (nie wolno korzystać z komputera i z książek).
- Za zaliczenie wykładów można uzyskać od 0 do 12 punktów.
- Ostateczna ocena z przedmiotu ustalana jest na podstawie sumy punktów z zaliczenia wykładów i laboratorium wg. następującej skali: 18-21,5 pkt. = 3.0, 22-24,5 pkt. = 3.5, 25-28,5 pkt. = 4.0, 29-31,5 pkt. = 4.5, 32-36pkt. = 5.0
- Zaliczenie przedmiotu po upływie regulaminowego terminu (ostatnim dniu zajęć semestru letniego) jest możliwe jedynie podczas testu, który odbędzie się w styczniu następnego roku kalendarzowego.

Literatura

- Tao Pang, *Metody obliczeniowe w fizyce*, PWN 2001
- Z. Fortuna, B. Macukow, J. Wąsowski, *Metody numeryczne*, WNT 2001
- A. Ralston, *Wstęp do analizy numerycznej*, PWN
- www.google.com

O czym mówić nie będziemy

- Błędy
- Oszacowania
- Zapis
- Obliczenia
- Wszystko, co wiąże się z rachunkiem błędów

Metody numeryczne – metody rozwiązywania problemów matematycznych za pomocą operacji na liczbach. Otrzymywane tą drogą wyniki są na ogół przybliżone, jednak dokładność obliczeń może być z góry określona i dobiera się ją zależnie od potrzeb.

Metody numeryczne wykorzystywane są wówczas gdy badany problem nie ma w ogóle rozwiązania analitycznego (danego wzorami), lub korzystanie z takich rozwiązań jest uciążliwe ze względu na ich złożoność.

W szczególności dotyczy to:

- całkowania
- znajdowania miejsc zerowych wielomianów
- rozwiązywania układów równań liniowych w przypadku większej liczby równań i niewiadomych
- rozwiązywania równań różniczkowych i układów takich równań
- znajdowania wartości i wektorów własnych
- aproksymacji, czyli przybliżaniu nieznanych funkcji

Obliczenia numeryczne a symboliczne

Obliczenia numeryczne: wykorzystują liczby bezpośrednio

(w celu uzyskania wyniku wykonują działania na liczbach)

Obliczenia symboliczne: liczby reprezentowane są przez symbole

(przekształcają symbole zgodnie z matematycznymi regułami by uzyskać symboliczny wynik)

Przykład (numeryczny)

$$\frac{(17.36)^2 - 1}{17.36 + 1} = 16.36$$

Przykład (symboliczny)

$$\frac{x^2 - 1}{x+1} = x-1$$

Rozwiązania analityczne a numeryczne

Rozwiązanie analityczne:

Dokładny wynik numeryczny lub symboliczny (może wykorzystywać symbole, np. tan(83), π, e).

Rozwiązanie numeryczne:

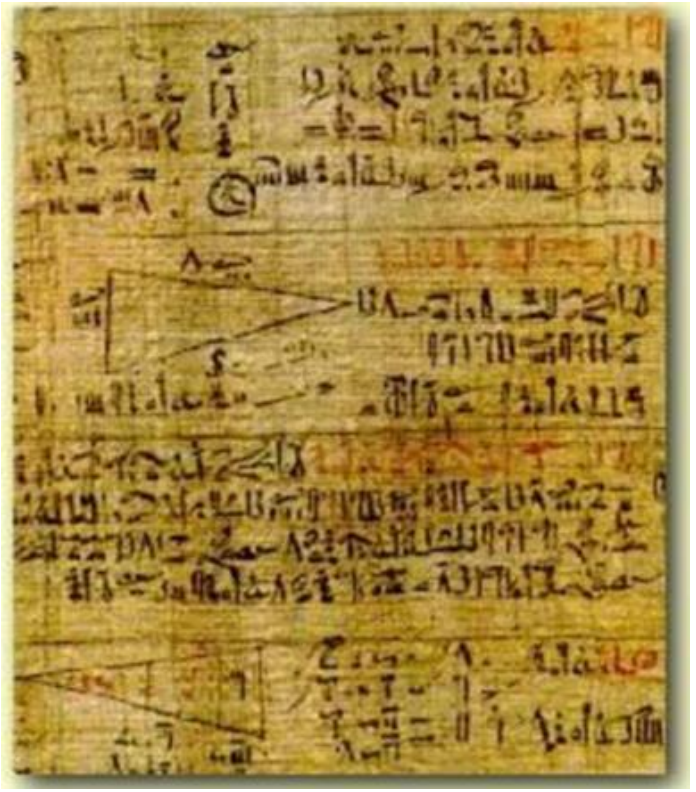
Wynik przedstawiony całkowicie numerycznie (niekoniecznie dokładny)

Przykład (analityczny)	Przykład (numeryczny)
$\frac{1}{4}$	0.25
$\frac{1}{3}$	0.33333 . . . (?)
π	3.14159 . . . (?)
tan (83)	0.88472 . . . (?)

Odrobina historii…

Papirus Rhinda (1650 p.n.e.)

egipski podręcznik arytmetyki i geometrii. Dowód posiadania przez Egipcjan szerszej wiedzy matematycznej, w szczególności znajomość liczb pierwszych, liczb złożonych, średnich arytmetycznej, geometrycznej i harmonicznej i uproszczonej wersji sita Eratostenesa. Sugeruje również znajomość pierwocin geometrii analitycznej. Znajdują się w nim bowiem:



metoda obliczenia liczby π z dokładnością lepszą niż 1%,
próba kwadratury koła
najstarsze znane użycie kotangensa.

Jednym z najdłuższych wątków w historii metod numerycznych była próba obliczenia liczby π z jak największą dokładnością. We wspomnianym papirusie wartość liczby π, przybliżano wartością

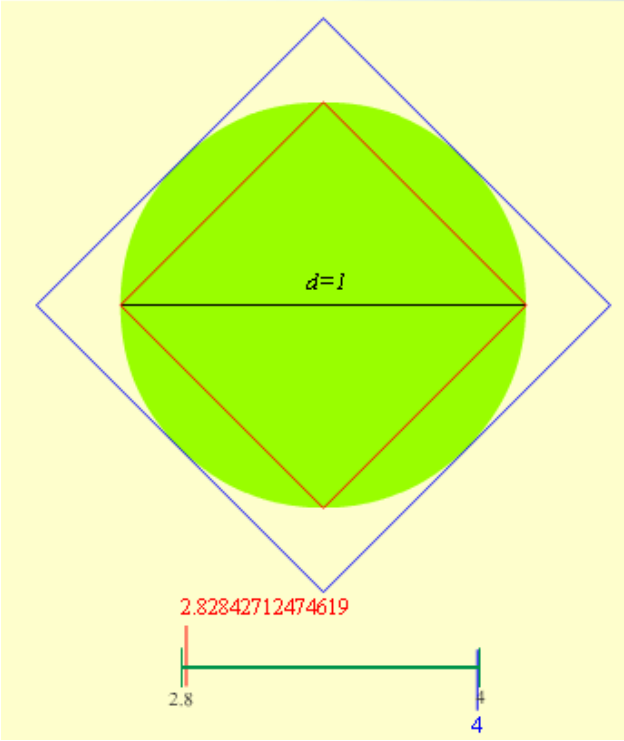
$$\frac{4^4}{3^4} = 3,1604\ldots$$

Archimedes 287-212 p.n.e.

Archimedes użył metody bazującej na zależnościach geometrycznych, pozwalającej oszacowywać π z (teoretycznie) dowolną dokładnością.

Algorytm Archimedesa

- znajdź długość obwodu wielokąta wpisanego w okrąg o promieniu 1/2
- znajdź długość obwodu wielokąta opisanego na okręgu o promieniu 1/2
- wartość π leży między tymi dwoma liczbami



Metoda Machina

Korzystając z faktu, że $\pi = 4 \arctan 1$

oraz $\arctan(x) = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \frac{x^9}{9} - \frac{x^{11}}{11} + \ldots$

około 1700 r. John Machin odkrył zależność

$$\pi = 16 \arctan\left(\frac{1}{5}\right) - 4 \arctan\left(\frac{1}{239}\right)$$

I jako pierwszy człowiek obliczył π z dokładnością do 100 cyfr

Inne przybliżenie: $\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \frac{1}{11} + \ldots.$

Użyte w 1973 roku do znalezienia pierwszego miliona cyfr.

Metody numeryczne – po co?

Dobrze dobrane metody numeryczne umożliwiają (ułatwiają) symulację zjawisk rzeczywistych

Przykłady katastrof związanych ze złymi obliczeniami numerycznymi:

I Eksplozja wartej 500 milionów $ rakiety Ariane 5 w 30 sekund po starcie z kosmodromu w Gujanie Francuskiej 4.06.1996

przyczyna: błędna konwersja 64-bitowej liczby zmiennoprzecinkowej na 16-bitową liczbę całkowitą (overflow):

```
y := int(x)
```

Metody numeryczne – po co?

Dobrze dobrane metody numeryczne umożliwiają (ułatwiają) symulację zjawisk rzeczywistych

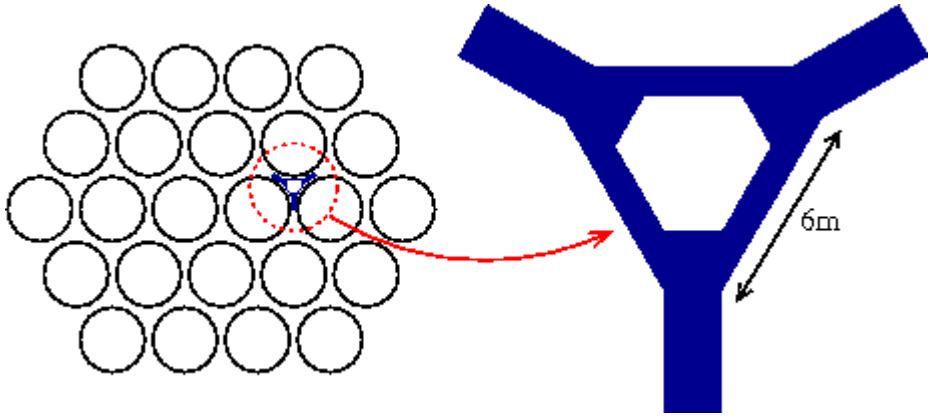
Przykłady katastrof związanych ze złymi obliczeniami numerycznymi:

I Eksplozja wartej 500 milionów $ rakiety Ariane 5 w 30 sekund po starcie z kosmodromu w Gujanie Francuskiej 4.06.1996

przyczyna: błędna konwersja 64-bitowej liczby zmiennoprzecinkowej na 16-bitową liczbę całkowitą (overflow):

```
y := int(x)
```

II Zatonięcie platformy wiertniczej Sleipner A na Morzu Północnym 23.08.1991 (1 miliard $) – przyczyna: niedokładność zamodelowania elementu konstrukcji za pomocą metody elementów skończonych.



III Tragedia w Dharan (Arabia Saudyjska, 21.02.1991) – 28 ofiar - błąd zaokrągleń w zegarze systemowym komputera.

Błąd w pomiarze czasu, który w momencie zaplanowanego strzału wynosił 1/3 s.

Błąd w określeniu pozycji wynoszący 687 metrów.

System Patriot konstruowany był z założeniem, że w celu utrudnienia lokalizacji nie będzie działał w jednym miejscu dłużej niż 8 godzin.

Armia izraelska zidentyfikowała ten błąd oprogramowania już przed feralnym dniem i 11 lutego 1991 r. poinformowała o nim producenta oprogramowania, jednakże odpowiednia "łata" dotarła do Arabii Saudyjskiej dopiero 22 lutego, a więc jeden dzień po uderzeniu Scuda w Dhahran.

Zapis komputerowy liczb

Liczby całkowite

Decimal	Conversion	Base 2
1	2^0	0000 0001
2	2^1	0000 0010
4	2^2	0000 0100
8	2^3	0000 1000
27	$2^4 + 2^3 + 2^1 + 2^0$	0000 0001

Zwykle zapis 16 lub 32 bitowy

Algorytm do konwersji liczb z systemu dziesiątkowego do binarnego

$$(N)_{10} = (b_j b_{j-1} \ldots b_2 b_0)_2$$
$$= b_j \cdot 2^j + \cdots + b_1 \cdot 2^1 + b_0 \cdot 2^0$$

Obliczmy $(N)_{10}/2 = Q + R$:

$$\frac{N}{2} = \underbrace{b_j \cdot 2^{j-1} + \cdots + b_1 \cdot 2^0}_{=Q} + \underbrace{\frac{b_0}{2}}_{=R}$$

Przykład konwersji: Dziesiętna liczba 11 w systemie dwójkowym

$$11/2 = 5R1 \quad \Rightarrow \quad b_0 = 1$$
$$5/2 = 2R1 \quad \Rightarrow \quad b_1 = 1$$
$$2/2 = 1R0 \quad \Rightarrow \quad b_2 = 0$$
$$1/2 = 0R1 \quad \Rightarrow \quad b_3 = 1$$

$11_{(10)} = 1011_{(2)}$

Liczby zmiennoprzecinkowe

Różne sposoby zapisu

- 123.456
- 123.456×10^0
- 1.23456×10^2

Weźmy $(12.52)_{10}$:

$$(12.52)_{10} = 1 \cdot 10^1 + 2 \cdot 10^0 + 5 \cdot 10^{-1} + 2 \cdot 10^{-2}$$

Analogicznie: $(1011.0011)_2$?

$$\begin{aligned}(101.0011)_2 &= 1 \cdot 2^2 + 0 \cdot 2^1 + 1 \cdot 2^0 + 0 \cdot 2^{-1} + 0 \cdot 2^{-2} + 1 \cdot 2^{-3} + 1 \cdot 2^{-4} \\ &= 4 + 1 + 1/8 + 1/6 \\ &= 5\frac{3}{16} \\ &= (5.1875)_{10}\end{aligned}$$

Algorytm do obliczania części ułamkowej *P* liczby binarnej

$$P = 0.b_{-1}b_{-2}b_{-3}\ldots$$
$$= b_{-1}\cdot 2^{-1} + \ldots$$

Pomnóżmy *P* przez *2*. Część całkowita *2P* to b_{-1}.

$$2P = b_{-1}\cdot 2^{0} + b_{-2}\cdot 2^{-1} + b_{-3}\cdot 2^{-2} + \ldots$$

Przykład

Obliczmy 0.625:

$$\begin{aligned} 2\cdot 0.625 = 1.25 &\Rightarrow b_{-1} = 1 \\ 2\cdot 0.25 = 0.5 &\Rightarrow b_{-2} = 0 \\ 2\cdot 0.5 = 1.0 &\Rightarrow b_{-3} = 1 \end{aligned}$$

$0.625_{(10)} = 0.101_{(2)}$

Zapis liczb zmiennoprzecinkowych

W systemie dziesiętnym:

$$\pm 0.d_1 d_2 d_3 \ldots d_n \times 10^s$$

$d_1d_2d_3$… - mantysa
s – wykładnik (cecha)

W systemie dwójkowym:

$$\pm 0.b_1 b_2 b_3 \ldots b_n \times \mathbf{2}^t$$

Przykład

32 bity

$$\underbrace{b}\ \underbrace{bb\ldots bbb}\ \underbrace{bbbbbbbb}$$

Znak mantysa (23 bity) cecha (8 bitów)

Przykład

64 bity

$$\underbrace{b}\ \underbrace{bb\ldots bbb}\ \underbrace{bbbbbbbbbbb}$$

Znak mantysa (52 bity) cecha (11 bitów)

typ	zakres
integer	$-32768 \leq I \leq 32767$
single	$-3.40 \times 10^{38} \leq x \leq -1.18 \times 10^{-38}$
	0
	$1.18 \times 10^{-38} \leq x \leq 3.40 \times 10^{38}$
double	$-1.80 \times 10^{318} \leq x \leq -2.23 \times 10^{-308}$
	0
	$1.80 \times 10^{-308} \leq x \leq 1.80 \times 10^{308}$

Zadanie

Przedstaw w postaci 24 bitowego kodu binarnego liczbę $0.1_{(10)}$

0.1	
0.2	0
0.4	0
0.8	0
1.6	1
1.2	1
0.4	0
0.8	0
1.6	1
1.2	1
0.4	0
0.8	0
1.6	1
1.2	1

0.1(10) = 0.00011001100110011001100110011001100110011…….

Mały błąd dokładności liczby 0.1
(0.00000000000000000000000011001100... dwójkowo, lub 0.000000095 dziesiątkowo)

Pomnóżmy to przez liczbę dziesiątych sekundy w ciągu 100 godzin

$$0.000000095 \times 100 \times 60 \times 60 \times 10 = 0.34$$

Scud leci z prędkością około 1676 metrów na sekundę, a zatem przebywa ponad pół kilometra w tym czasie - wystarczająco dużo, by znaleźć się poza zasięgiem rakiety Patriot…

Zagadka nr 1

```
double funkcja()
{
    double x=0;
    while(x != 1)
    {
      x += 0.1;
    }
    return x;
}
```

Zagadka nr 2

```
double A,B,C,x;
A=1;
B=1E9;
C=1;
x=(-B+sqrt(B*B-4*A*C))/(2*A);
```

Dodawanie a + b spowoduje duży błąd, gdy

- $a \gg b$
- $a \ll b$

Niech:

$$a = x.xxx \cdots \times 10^{0}$$
$$b = y.yyy \cdots \times 10^{-8}$$

Wtedy:

skończona precyzja

$$\begin{array}{rll} & x.xxx\ xxxx\ xxxx\ xxxx & \\ + & 0.000\,0000\ yyyy\ yyyy & yyyy\ yyyy \\ \hline = & x.xxx\ xxxx\ zzzz\ zzzz & \underbrace{????\ ????}_{\text{utracona precyzja}} \end{array}$$

Prawa algebry nie obowiązują

Przemienność dodawania: $a + (b + c) = (a + b) + c$

$a = 0.123\,41 \times 10^5 \quad b = -0.123\,40 \times 10^5 \quad c = 0.143\,21 \times 10^1$

$a + (b + c)$

$= 0.123\,41 \times 10^5 + (-0.123\,40 \times 10^5 + 0.143\,21 \times 10^1)$

$= 0.123\,41 \times 10^5 - 0.123\,39 \times 10^5$

$= 0.200\,00 \times 10^1$

$(a + b) + c$

$= (0.123\,41 \times 10^5 - 0.123\,40 \times 10^5) + 0.143\,21 \times 10^1$

$= 0.100\,00 \times 10^1 + 0.143\,21 \times 10^1$

$= 0.243\,21 \times 10^1$

Wyniki różnią się o ponad 20%!

Jak obliczyć precyzję obliczeń?

```
float epsilon=1, b;
int it = 1;
while(1)
   {
   epsilon = epsilon / 2;
   b = 1 + epsilon;
   if(b==1) break;
   it++;
   };
```

Uwagi praktyczne

- Porównując liczby zmiennoprzecinkowe nie używajcie bitowych operatorów równości (czyli ==)

- Skończcie iteracje, gdy osiągniecie precyzję obliczeń na liczbach zmiennoprzecinkowych

ŹLE

```
if(x==y)
.
.
.
end
```

DOBRZE

```
if(abs(x-y) < tol)
.
.
.
end
```

Pentium bug

Rozważmy równość:

$$A - \frac{A}{B} * B = 0$$

Zero czy nie zero?

- $A = 4,295,835.0$ and $B = 3,145,727.0$
- $A - \frac{A}{B} * B = 256$ on a Pentium

Q: Ilu inżynierów Intela potrzeba do wkręcenia jednej żarówki?

A: 0.999999325678