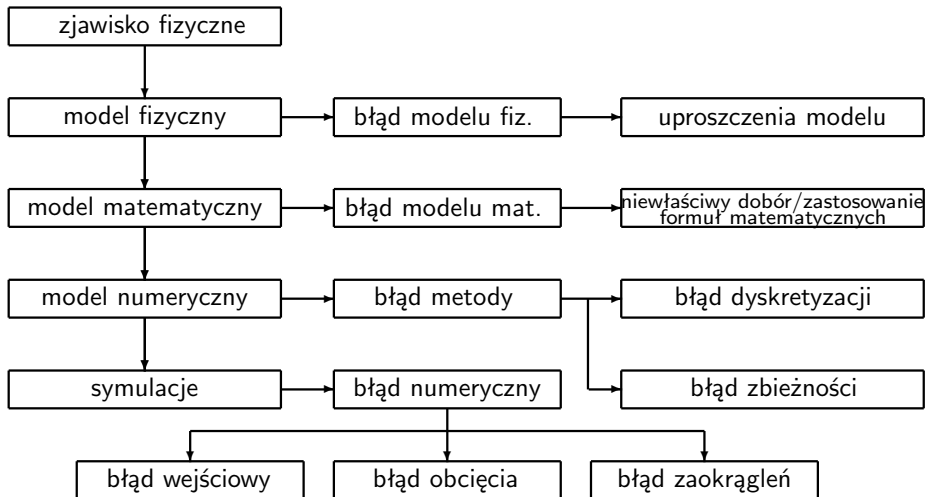


Metody numeryczne

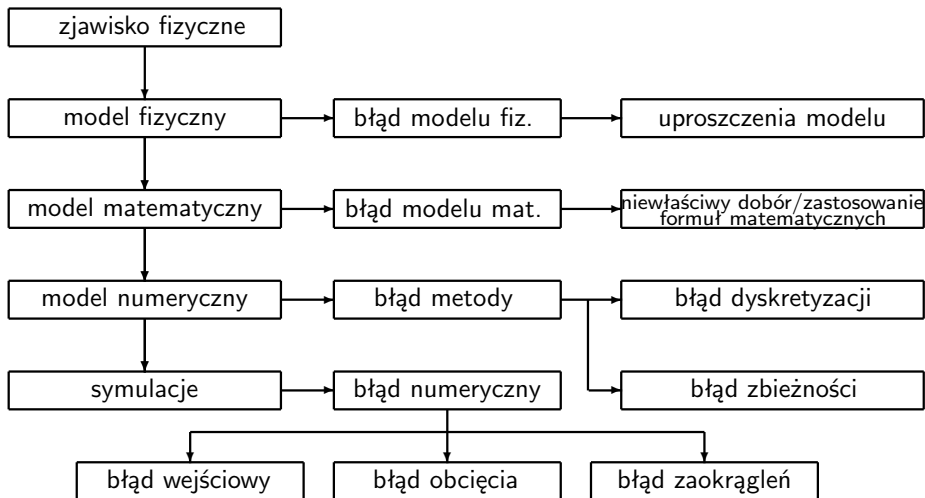
Wykład 1

Zagadnienia wstępne, arytmetyka zmiennoprzecinkowa

Błędy w modelowaniu numerycznym



Błędy w modelowaniu numerycznym



$$\sin(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots$$

Błąd jako miara dokładności

x – wartość dokładna

$\tilde{x}$ – wartość przybliżona

znane x	nieznane x
błąd bezwzględny	kres górny b.b.
$\varepsilon_x = \tilde{x} - x $	$E_x > \varepsilon_x$
błąd względny	kres górny b.w.
$\delta_x = \frac{\varepsilon_x}{x}, \quad x \neq 0$	$\Delta_x = \frac{E_x}{\tilde{x}}, \quad \tilde{x} \neq 0$

Przykład

$$x = \pi = 3.14159265358979\dots$$

$$\tilde{x} = 3.14$$

$$\varepsilon_x = ?, \delta_x = ?$$

$$\varepsilon_x = |\pi - 3.14| \leq |3.14159265358979\dots - 3.14| = 0.00159265358979\dots$$

$$\Rightarrow E_x = 0.0016$$

$$\begin{aligned} \Delta_x &= \frac{0.0016}{3.14} \approx 0.00050955414013 = \\ &= 5.0955414013 \cdot 10^{-4} \approx 0.05096\% \end{aligned}$$

Reprezentacja zmiennopozycyjna

$$x = (-1)^S \cdot M \cdot B^E$$

S – znak, $S = \{0, 1\}$

M – mantysa, znormalizowana: $M \in [1, B)$; rzadziej $M \in [\frac{1}{B}, 1)$

B – podstawa systemu liczbowego

E – cecha,

Przykład

$$-987.6543 = (-1)^1 \cdot 9.876543 \cdot 10^2$$

$$0.0010101 = (-1)^0 \cdot 1.0101 \cdot 2^{-3}$$

$$-0.21102 = (-1)^1 \cdot 2.1102 \cdot 3^{-1}$$

$$A1F73.02 = (-1)^0 \cdot A.1F7302 \cdot 16^4$$

System binarny - konwersja

- Liczby całkowite i stałoprzecinkowe $\rightarrow$ kodowanie liczb ujemnych: znak moduł (ZM), kod uzupełnień do 1 (U1), do 2 (U2)
- Konwersja bin $\rightarrow$ dec / dec $\rightarrow$ bin

Przykład

$$(11001.01)_2 = 1 \cdot 2^4 + 1 \cdot 2^3 + 0 \cdot 2^2 + 0 \cdot 2^1 + 1 \cdot 2^0 + 0 \cdot 2^{-1} + 1 \cdot 2^{-2} = 16 + 8 + 1 + 0.25 = (25.25)_{10}$$

Przykład

$$(23.1)_{10} = (?)_2$$

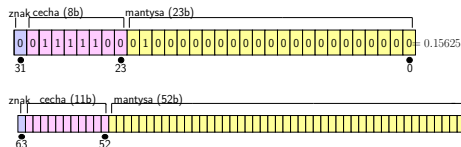
$\div 2$	r		$\times 2$	
23	1	0.1	0.2	
11	1	0.2	0.4	
5	1	0.4	0.8	
2	0	0.8	1.6	
1	1	0.6	1.2	
0		0.2	0.4	
		$\vdots$	$\vdots$	
		$\vdots$	$\vdots$	

$= (10111.000110011\dots)_2$

- Stała pozycja przecinka $\rightarrow$ ograniczony zakres

Binarne systemy zmiennoprzecinkowe – standard IEEE-754

$$x = (-1)^S \cdot M \cdot 2^{E-bias}$$



Typ IEEE-754	Pojedynczej precyzji	Podwójnej precyzji
Nazwa typu w C, Java	float	double
Nazwa typu w Pascalu	single	real lub double
Nazwa typu w Fortranie	real*4	real*8
Bit znaku	1	1
Liczba bitów cechy	8	11
Liczba bitów mantysy	23	52
Rozmiar (w bitach)	32	64
Bias	127	1023
Orientacyjny zakres	$10^{-38} \dots 10^{38}$	$10^{-308} \dots 10^{308}$
Orientacyjna precyzja	$6 * 10^{-8}$	10^{-16}

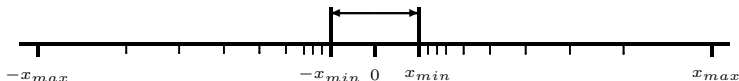
Matlab/Octave - domyślny typ: double

Orientacyjna precyzja == epsilon maszynowe

Zakres liczb reprezentowanych w danym zapisie:

$$[-x_{max}, -x_{min}] \cup \{0\} \cup [x_{min}, x_{max}]$$

Zero maszynowe



Błąd względny reprezentacji: $\Delta_x = \frac{1}{B^{m-1}}$

$|x| < x_{min} \rightarrow$ błąd niedomiaru (underflow)

$|x| > x_{max} \rightarrow$ przepełnienie/błąd nadmiaru (overflow) $\rightarrow$ przerwanie obliczeń

Przykład

Przedstaw wszystkie liczby maszynowe zapisane na 5 bitach. Przyjąć 2 bity na cechę, 2 na mantysę, bias = 1.

Przykład

Przedstaw wszystkie liczby maszynowe zapisane na 5 bitach. Przyjąć 2 bity na cechę, 2 na mantysę, bias = 1.

S	C	C	M	M
---	---	---	---	---

$$x = (-1)^S \cdot M \cdot 2^{E-bias}$$

cecha:

0 0	→	(0) ₁₀
0 1	→	(1) ₁₀
1 0	→	(2) ₁₀
1 1	→	(3) ₁₀

$$E \in \{0, 1, 2, 3\} \rightarrow$$

$$E - bias \in \{-1, 0, 1, 2\}$$

mantysa:

1. 0 0	→	A
1. 0 1	→	B
1. 1 0	→	C
1. 1 1	→	D

	-1	0	1	2
A	0.100	1.00	10.0	100
B	0.101	1.01	10.1	101
C	0.110	1.10	11.0	110
D	0.111	1.11	11.1	111

	-1	0	1	2
A	0.5	1	2	4
B	0.625	1.25	2.5	5
C	0.75	1.5	3	6
D	0.875	1.75	3.5	7

Skracanie: ucinanie i zaokrąglanie

Ucinanie – odrzucenie cyfr na prawo od t -tej

- wada: wprowadzany błąd ma znak systematycznie przeciwny niż skracana liczba

$$- \varepsilon_x < 10^{-t}$$

Zaokrąglanie - wybór takiej liczby o t cyfrach ułamkowych, która jest najbliższa do danej:

- jeśli $t + 1$ cyfra $\leq 4 \rightarrow$ ucinanie
- jeśli $t + 1$ cyfra ≥ 5 i $\leq 9 \rightarrow$ do cyfry t -tej należy dodać 1
- jeśli $t + 1$ cyfra = 5
 - jeśli wśród odrzuconych cyfr są niezera do t -tej dodaje się jeden
 - jeśli odrzucane cyfry są zerami, to do t -tej cyfry:
 - dodaje się 1 jeśli jest nieparzysta
 - pozostawia bez zmian jeśli parzysta

$$- \varepsilon_x < \frac{1}{2} \cdot 10^{-t}$$

Przykład

x	round(x)	chop(x)
0.2392	0.239	0.239
0.2397	0.240	0.239
-0.2397	-0.240	-0.239
0.23652	0.237	0.236
0.23750	0.238	0.237
0.23650	0.236	0.236

Błąd zaokrągleń

Arytmetyka zmiennoprzecinkowa

$$(x + y) + z \neq x + (y + z)$$

$$(x \cdot y) \cdot z \neq x \cdot (y \cdot z)$$

$$x \cdot (y + z) \neq x \cdot y + x \cdot z$$

$$x_1 = (-1)^{S_1} \cdot M_1 \cdot B^{E_1}$$

$$x_2 = (-1)^{S_2} \cdot M_2 \cdot B^{E_2}$$

$$x_1 \pm x_2 = (M_1 \pm M_2 \cdot B^{(E_2 - E_1)}) \cdot B^{E_1} \leftarrow \text{denormalizacja liczby o mniejszym wykładniku}$$

→ błędy obcięcia ($E_1 - E_2 > \text{liczba bitów mantysy} !!!$)

$$x_1 \cdot x_2 = (S_1 \oplus S_2)(M_1 \cdot M_2) \cdot B^{E_1 + E_2}$$

$$x_1 / x_2 = (S_1 \oplus S_2)(M_1 / M_2) \cdot B^{E_1 - E_2}$$

gdzie $\oplus$ – exor

Standard IEEE-754 cd

Wartość specjalna	Bit znaku	Bity wykładnika	Bity mantysy	Uwagi
NaN	x	111..111	xxxx...xxx	wszystkie bity wykładnika = 1, mantysa $\neq 0$
QNaN	x	111..111	1xxx...xxx	uwagi jak dla NaN, pierwszy bit mantysy zawsze = 1
SNaN	x	111..111	0xxx...xxx	uwagi jak dla NaN, pierwszy bit mantysy zawsze = 0
Zero	x	000..000	0000...000	wszystkie bity mantysy i wykładnika = 0, bit znaku decyduje o znaku wartości
$\pm$ Nieskończoność	x	111..111	0000...000	wszystkie bity mantysy = 0, wszystkie bity wykładnika = 1, bit znaku decyduje o znaku wartości
Nieznormalizowana	x	000..000	xxxx...xxx	mantysa $\neq 0$, wszystkie bity wykładnika = 0, bit znaku decyduje o znaku wartości

- Zera: $+0$ ($[00000000]_{16}$), -0 ($[80000000]_{16}$).
 -0 jeśli nastąpił niedomiar, $+0$ w przeciwnym przypadku.
- Nieskończoności: $+\infty$ jeśli wystąpił nadmiar, $x + \infty = +\infty$,
 $x \cdot \infty = +\infty$, $\infty/x = +\infty$ dla $x > 0$.
- NaN "not a number" ($e = 255$, $f \neq 0$).
 NaN jeśli $0/0$, $\infty - \infty$, $\text{NaN} + x$.

liczby denormalizowane – zapętlienie
"próżni" wokół zera ($-x_{min}, x_{min}$)

Teoretycznie dla

$$x < 2^{1-1023} \approx 2.2251^{-308}, x \rightarrow 0$$

W praktyce:

[illegible]

Przenoszenie się błędów

Przykład

Niech $x_1 = \tilde{x}_1 \pm \varepsilon_1$, $x_2 = \tilde{x}_2 \pm \varepsilon_2$

Ile wyniesie błąd oszacowań $x_1 + x_2$, $x_1 - x_2$, $x_1 x_2$, x_1/x_2 .

Przenoszenie się błędów

Przykład

Niech $x_1 = \tilde{x}_1 \pm \varepsilon_1$, $x_2 = \tilde{x}_2 \pm \varepsilon_2$ ($\implies x_1 \in \langle \tilde{x}_1 - \varepsilon_1, \tilde{x}_1 + \varepsilon_1 \rangle, \dots$)
Ile wyniesie błąd oszacowań $x_1 + x_2$, $x_1 - x_2$, $x_1 x_2$, x_1/x_2 .

Przenoszenie się błędów

Przykład

Niech $x_1 = \tilde{x}_1 \pm \varepsilon_1$, $x_2 = \tilde{x}_2 \pm \varepsilon_2$

Ile wyniesie błąd oszacowań $x_1 + x_2$, $x_1 - x_2$, $x_1 x_2$, x_1/x_2 .

$$\tilde{x}_1 - \varepsilon_1 + (\tilde{x}_2 - \varepsilon_2) \leq x_1 + x_2 \leq \tilde{x}_1 + \varepsilon_1 + (\tilde{x}_2 + \varepsilon_2)$$

$$\tilde{x}_1 + \tilde{x}_2 - (\varepsilon_1 + \varepsilon_2) \leq x_1 + x_2 \leq \tilde{x}_1 + \tilde{x}_2 + (\varepsilon_1 + \varepsilon_2)$$

$$x_1 + x_2 = \tilde{x}_1 + \tilde{x}_2 \pm (\varepsilon_1 + \varepsilon_2)$$

Przenoszenie się błędów

Przykład

Niech $x_1 = \tilde{x}_1 \pm \varepsilon_1$, $x_2 = \tilde{x}_2 \pm \varepsilon_2$

Ile wyniesie błąd oszacowań $x_1 + x_2$, $x_1 - x_2$, $x_1 x_2$, x_1/x_2 .

$$\tilde{x}_1 - \varepsilon_1 + (\tilde{x}_2 - \varepsilon_2) \leq x_1 + x_2 \leq \tilde{x}_1 + \varepsilon_1 + (\tilde{x}_2 + \varepsilon_2)$$

$$\tilde{x}_1 + \tilde{x}_2 - (\varepsilon_1 + \varepsilon_2) \leq x_1 + x_2 \leq \tilde{x}_1 + \tilde{x}_2 + (\varepsilon_1 + \varepsilon_2)$$

$$x_1 + x_2 = \tilde{x}_1 + \tilde{x}_2 \pm (\varepsilon_1 + \varepsilon_2)$$

Analogicznie

$$\tilde{x}_1 - \varepsilon_1 - (\tilde{x}_2 + \varepsilon_2) \leq x_1 - x_2 \leq \tilde{x}_1 + \varepsilon_1 - (\tilde{x}_2 - \varepsilon_2)$$

$$\tilde{x}_1 - \tilde{x}_2 - (\varepsilon_1 + \varepsilon_2) \leq x_1 - x_2 \leq \tilde{x}_1 - \tilde{x}_2 + (\varepsilon_1 + \varepsilon_2)$$

A więc:

$$x_1 - x_2 = \tilde{x}_1 - \tilde{x}_2 \pm (\varepsilon_1 + \varepsilon_2)$$

→ arytmetyka (analiza) przedziałowa

→ algebra liczb rozmytych

Utrata cyfr znaczących

Niech $x = 0.2345 \pm 0.00005$ i $y = 0.2344 \pm 0.00005$
(tzn. $x \in \langle 0.23445, 0.23455 \rangle$, $y \in \langle 0.23435, 0.23445 \rangle$)
a więc ($x \approx y$).

$$x - y = 0.0001 \pm 0.0001$$

$$\delta_{x-y} = \frac{\varepsilon_{x-y}}{x-y} = 100\%!$$

Przenoszenie się błędów/Szacowanie błędów funkcji

Przykład

Korzystając z definicji błędu względnego

$$(\delta_x = \frac{\varepsilon_x}{x} = \frac{|\tilde{x}-x|}{x} \implies \tilde{x} = x \pm x\delta_x = x(1 \pm \delta_x)$$

$$\tilde{x}_1 = x_1(1 \pm \delta_1)$$

$$\tilde{x}_2 = x_2(1 \pm \delta_2)$$

$$\begin{aligned}\tilde{x}_1\tilde{x}_2 &= x_1(1 \pm \delta_1)x_2(1 \pm \delta_2) = x_1x_2(1 \pm \delta_1)(1 \pm \delta_2) = \\ &= x_1x_2(1 \pm \delta_1 \pm \delta_2 \pm \delta_1\delta_2)\end{aligned}$$

$$\delta_1 \ll 1, \delta_2 \ll 1 \implies \delta_1\delta_2 \approx 0$$

A więc

$$\delta_{x_1x_2} \approx \pm(\delta_1 + \delta_2)$$

Przenoszenie się błędów/Szacowanie błędów funkcji

Przykład

Korzystając z definicji błędu względnego

$$(\delta_x = \frac{\varepsilon_x}{x} = \frac{|\tilde{x}-x|}{x} \implies \tilde{x} = x \pm x\delta_x = x(1 \pm \delta_x)$$

$$\tilde{x}_1 = x_1(1 \pm \delta_1)$$

$$\tilde{x}_2 = x_2(1 \pm \delta_2)$$

$$\tilde{x}_1\tilde{x}_2 = x_1(1 \pm \delta_1)x_2(1 \pm \delta_2) = x_1x_2(1 \pm \delta_1)(1 \pm \delta_2) = x_1x_2(1 \pm \delta_1 \pm \delta_2 \pm \delta_1\delta_2)$$

$$\delta_1 \ll 1, \delta_2 \ll 1 \implies \delta_1\delta_2 \approx 0$$

A więc

$$\delta_{x_1x_2} \approx \pm(\delta_1 + \delta_2)$$

Analogicznie:

$$\frac{\tilde{x}_1}{\tilde{x}_2} = \frac{x_1(1 \pm \delta_1)}{x_2(1 \pm \delta_2)} = \frac{x_1}{x_2} \frac{(1 \pm \delta_1)(1 \pm \delta_2)}{(1 \pm \delta_2)(1 \pm \delta_2)} = \frac{x_1}{x_2} \frac{(1 \pm \delta_1 \pm \delta_2 \pm \delta_1\delta_2)}{1 - (\pm\delta_2)^2} = \frac{x_1}{x_2} (1 \pm \delta_1 \pm \delta_2)$$

$$\delta_{x_1/x_2} \approx \pm(\delta_1 + \delta_2)$$

Niech y będzie funkcją zmiennych $x_1, x_2, \dots, x_n$, $y = y(x_1, x_2, \dots, x_n)$ i niech dane będą przybliżone wartości $\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n$ zmiennych $x_1, x_2, \dots, x_n$, których dokładnymi wartościami są $\hat{x}_1, \hat{x}_2, \dots, \hat{x}_n$.
 Przyjmując $\tilde{\mathbf{x}} = [\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n]$, oraz $\hat{\mathbf{x}} = [\hat{x}_1, \hat{x}_2, \dots, \hat{x}_n]$ można zapisać

$$\varepsilon_{x_i} = \tilde{x}_i - \hat{x}_i$$

$$\varepsilon_y = y(\tilde{\mathbf{x}}) - y(\hat{\mathbf{x}})$$

Twierdzenie

Ogólny wzór na przenoszenie się błędów:

$$\varepsilon_y = \sum_{i=1}^n \left| \frac{\partial y}{\partial x_i} \right| \cdot |\varepsilon_{x_i}|$$

Wzór określa błąd maksymalny, a więc uwzględnia najgorsze możliwe przypadki.

Przykład

$$y = x_1^2 - x_2$$

$$x_1 = 1.03 \pm 0.01$$

$$x_2 = 0.45 \pm 0.01$$

$$\varepsilon_y = ?$$

$$\frac{\partial y}{\partial x_1} = 2x_1 \quad \frac{\partial y}{\partial x_2} = -1$$

$$\varepsilon_y = \sum_{i=1}^2 \left| \frac{\partial y}{\partial x_i} \right| \cdot |\varepsilon_{x_i}| = |2x_1| |\varepsilon_{x_1}| + |-1| |\varepsilon_{x_2}| =$$

$$2 \cdot 1.03 \cdot 0.01 + 1 \cdot 0.01 = 0.0306$$

$$y = 1.03^2 - 0.45 \pm \varepsilon_y = 0.6109 \pm 0.0306$$

$$y \in < 0.6109 - 0.0306, 0.6109 + 0.0306 >$$

Uwarunkowanie zadania numerycznego

Definicja

Jeśli niewielkie względne zmiany danych powodują duże względne odkształcenia wyników, to zadanie nazywane jest źle uwarunkowanym. W przeciwnym przypadku mówi się o zadaniu dobrze uwarunkowanym

Przykład - funkcja 1D

Zbadać uwarunkowanie zadania obliczania wartości funkcji $f(x) \neq 0$ w punkcie x .

W tym celu "zaburzamy" dane x : $\tilde{x} = x(1 + \delta_x) = x + x\delta_x$. Oszacowanie względnej zmiany wyniku wyniesie:

$$\frac{|f(\tilde{x}) - f(x)|}{|f(x)|} = \frac{|f(x + x\delta_x) - f(x)|}{|f(x)|} \approx \frac{|f(x + x\delta_x) - f(x)|}{|f(x)|} \left| \frac{\delta_x}{\delta_x} \right| \approx \left| \frac{f'(x)}{f(x)} \delta_x \right| \leq$$

$$\left| \frac{x f'(x)}{f(x)} \right| \delta_x$$

↑
wskaźnik uwarunkowania zadania ($\rightarrow$ conditioning)

$|cond| \approx 0 \rightarrow$ dobrze uwarunkowane zadanie

$|cond| \gg 0 \rightarrow$ zadanie źle uwarunkowane

Twierdzenie

Wskaźnik uwarunkowania zadania obliczania wartości funkcji wielu zmiennych ($y = y(x_1, x_2, \dots, x_n)$) wynosi

$$\delta_y = \sum_{i=1}^n \left| \frac{x_i}{y} \frac{\partial y}{\partial x_i} \right| \cdot |\delta_{x_i}| \quad (1)$$

Przykład (EF)

Sprawdzić uwarunkowanie zadania obliczania wartości funkcji

$$y = y(p, q) = -p + \sqrt{p^2 + q}.$$

$$\frac{\partial y}{\partial p} = -1 + \frac{2p}{2\sqrt{p^2 + q}} \quad \frac{\partial y}{\partial q} = \frac{1}{2\sqrt{p^2 + q}}$$

$$\begin{aligned} \delta_y &= \sum_{i=1}^n \left| \frac{x_i}{y} \frac{\partial y}{\partial x_i} \right| \cdot |\delta_{x_i}| = \left| \frac{p}{y} \frac{\partial y}{\partial p} \right| \delta_p + \left| \frac{q}{y} \frac{\partial y}{\partial q} \right| \delta_q = \\ &= \left| \frac{p}{-p + \sqrt{p^2 + q}} \left(-1 + \frac{p}{\sqrt{p^2 + q}} \right) \right| \delta_p + \left| \frac{q}{-p + \sqrt{p^2 + q}} \left(\frac{1}{2\sqrt{p^2 + q}} \right) \right| \delta_q = \\ &= \left| \frac{p}{-p + \sqrt{p^2 + q}} \left(\frac{p - \sqrt{p^2 + q}}{\sqrt{p^2 + q}} \right) \right| \delta_p + \left| \frac{q(\sqrt{p^2 + q} + p)}{(p^2 + q - p^2)2\sqrt{p^2 + q}} \right| \delta_q = \\ &= \left| \frac{-p}{\sqrt{p^2 + q}} \right| \delta_p + \left| \frac{\sqrt{p^2 + q} + p}{2\sqrt{p^2 + q}} \right| \delta_q \end{aligned}$$

Przykład

$$y = y(p, q) = -p + \sqrt{p^2 + q}$$

$$\delta_y = \left| \frac{-p}{\sqrt{p^2 + q}} \right| \delta_p + \left| \frac{\sqrt{p^2 + q} + p}{2\sqrt{p^2 + q}} \right| \delta_q = c_p \delta_p + c_q \delta_q$$

Jak wpływają błędy δ_p i δ_q na δ_y

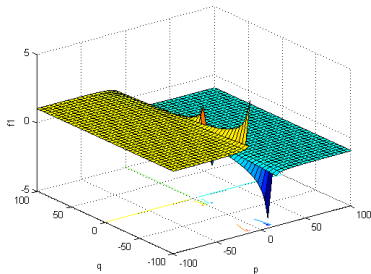
$p^2 + q > 0 \implies p \in R, q \in (-p^2, +\infty)$ – dziedzina funkcji

Dla $q > 0$: $c_p \approx 1$, $c_q \approx 1$ – zad. dobrze uwar.

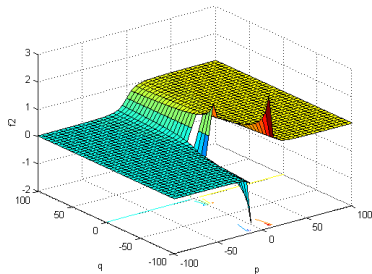
Dla $q < 0$:

$q \rightarrow -p^2 \implies \sqrt{p^2 + q} \rightarrow 0 \implies c_p \rightarrow \infty, c_q \rightarrow \infty$ – zad. źle uwar.

$$c_p = \left| \frac{-p}{\sqrt{p^2 + q}} \right|$$



$$c_q = \left| \frac{\sqrt{p^2 + q} + p}{2\sqrt{p^2 + q}} \right|$$



Zadanie

Korzystając ze wzoru (1) sprawdź uwarunkowanie dodawania i mnożenia dwóch liczb.

Stabilność numeryczna

Definicja

Algorytm jest numerycznie stabilny jeśli gwarantuje rozwiązania (dla dowolnego zadania rozważnej klasy) z błędem tego samego rzędu co optymalny poziom błędu rozwiązania danego zadania.

Prościej:

Błąd w algorytmie numerycznie stabilnym jest na poziomie „nieuniknionego” błędu wynikającego z przybliżonej reprezentacji danych i wyniku.

Jeszcze prościej:

Błąd w algorytmie niestabilnym popełniony na którymś etapie obliczeń rośnie w następnych etapach zniekształcając ostateczne wyniki.

Przykład

Rozważmy ciąg:

$$(a) \ x_0 = 1, x_1 = \frac{1}{3}, x_{n+1} = \frac{13}{3}x_n - \frac{4}{3}x_{n-1} \quad \Leftrightarrow \quad (b) \ x_n = \left(\frac{1}{3}\right)^n$$

	(b)	(a)
x_0	1.0000e + 00	1.0000e + 00
x_1	3.3333e - 01	3.3333e - 01
x_2	1.1111e - 01	1.1111e - 01
x_3	3.7037e - 02	3.7037e - 02
x_4	1.2346e - 02	1.2346e - 02
x_5	4.1152e - 03	4.1152e - 03
x_6	1.3717e - 03	1.3717e - 03
x_7	4.5725e - 04	4.5725e - 04
x_8	1.5242e - 04	1.5242e - 04
x_9	5.0805e - 05	5.0805e - 05
x_{10}	1.6935e - 05	1.6935e - 05
x_{11}	5.6450e - 06	5.6450e - 06
x_{12}	1.8817e - 06	1.8815e - 06
x_{13}	6.2723e - 07	6.2639e - 07
x_{14}	2.0908e - 07	2.0575e - 07
x_{15}	6.9692e - 08	5.6399e - 08
x_{16}	2.3231e - 08	-2.9941e - 08
x_{17}	7.7435e - 09	-2.0494e - 07
x_{18}	2.5812e - 09	-8.4816e - 07
x_{19}	8.6039e - 10	-3.4021e - 06

$$x_1 = (0.(3))_{10} \rightarrow (\dots)_2$$

$$e_1 = eps \approx 2.22 \cdot 10^{-16}$$

$$e_{n+1} = \frac{13}{3}e_n + \frac{4}{3}e_{n-1}$$

:

$$e_{16} > \left(\frac{13}{3}\right)^{15} e_1$$

$$e_{16} > 3.57 \cdot 10^9 \cdot 2.22 \cdot 10^{-16}$$

$$e_{16} > 10^{-7}$$

Kontrprzykład

Rozważmy ciąg:

$$x_0 = 1, x_1 = 4, x_{n+1} = \frac{13}{3}x_n - \frac{4}{3}x_{n-1} \quad \Leftrightarrow \quad x_n = (4)^n$$

Dziękuję za uwagę

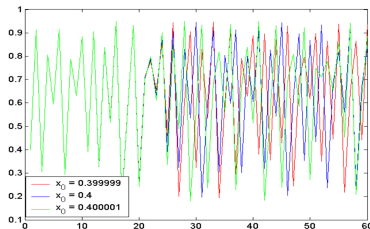
"Błąd" wejściowy – dynamika nieliniowa – przykład

- równanie logistyczne – Pierre Francois Verhulst (1825)
- wersja dyskretna – Robert May (1976)

$$x_{i+1} = \alpha x_i(1 - x_i)$$

α - współczynnik tempa wzrostu populacji

- $\alpha \in [3.57, 4]$ – zachowanie chaotyczne



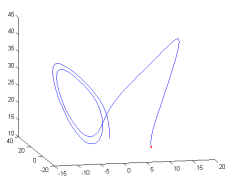
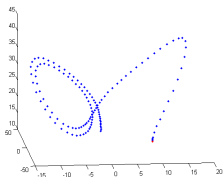
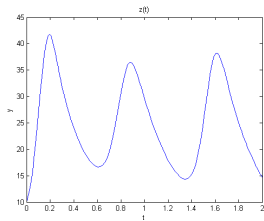
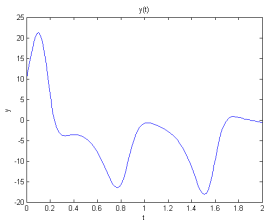
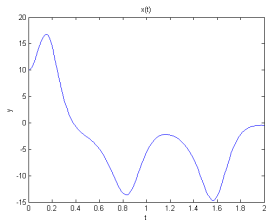
Powrót

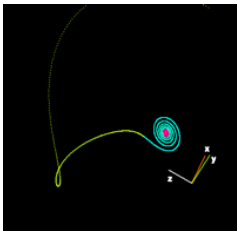
Układ Lorenza – konwekcja termiczna w atmosferze

$$\begin{aligned}\frac{dx}{dt} &= \sigma(y - x), \\ \frac{dy}{dt} &= x(\rho - z) - y, \\ \frac{dz}{dt} &= xy - \beta z.\end{aligned}\tag{2}$$

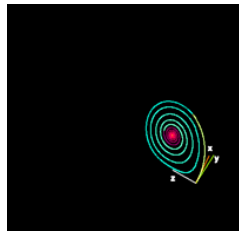
σ – stała charakteryzująca lepkość ośrodka, ρ – stała charakteryzująca przewodnictwo cieplne ośrodka, β – stała charakteryzująca rozmiary obszaru przepływu konwekcyjnego
 x – ruch konwekcyjny, y – różnice temperatur, z – rozkład pionowy temperatury

Powrót

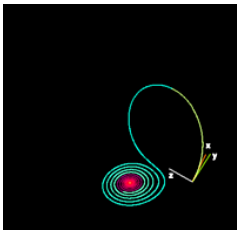




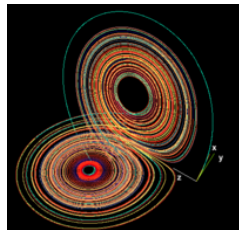
$$\rho = 14, \sigma = 10, \beta = \frac{8}{3}$$



$$\rho = 13, \sigma = 10, \beta = \frac{8}{3}$$



$$\rho = 15, \sigma = 10, \beta = \frac{8}{3}$$



$$\rho = 28, \sigma = 10, \beta = \frac{8}{3}$$

<http://www.youtube.com/watch?v=FYE4JKAXSFY>

Powrót

"Błąd" zaokrąglenia – przykład

- 25.02.1991, Dharan, Arabia Saudyjska – śmierć 28 amer. żołnierzy
- zegar 1/10 s: $(\frac{1}{10})_{10} = (0.0(0011))_2$
- dla 24b rejestru błąd dla jednego impulsu zegarowego
$$E = 0.0(0011) - 0.00011001100110011001100 = (0.00000000000000000000000011001100...)_{2} \approx (9.5 \cdot 10^{-8})_{10}[s]$$
- błąd dla 100h pracy systemu (vs 8h)
$$E_{tot} = 100 \cdot 60 \cdot 60 \cdot 10 \cdot 9.5 \cdot 10^{-8} \approx 0.3333s$$
- efekt błędu komputerowego przy nieprzerwanym działaniu systemu

Godziny	Sekundy	Sek. obliczone	Błąd w [s]	Błąd w [m]
0	0	0	0	0
1	3600	3599.9966	0.0034	7
8	28800	28799.9725	0.0275	55
20	72000	71999.9313	0.0687	137
48	172800	172799.352	0.1648	330
72	259200	259199.7528	0.2472	494
100	360000	359999.6667	0.3333	687

Źródło: <http://fas.org/spp/starwars/gao/im92026.htm>

- Rozwiązanie: a) reset systemu b) zliczanie impulsów np. co 0.125 s ($0.125_{10} = 0.001_2$)

Ograniczona liczba cyfr przeznaczonych na mantysę i cechę $\rightarrow$ skończona dokładność, ograniczony zbiór reprezentowanych wartości.

Przykład

Niech:

m – liczba cyfr przeznaczonych na mantysę

$n + 1$ – liczba cyfr przeznaczonych na wykładnik (n cyfr dla wartości i 1 dla znaku)

$$\begin{aligned}x &= (-1)^S \cdot M \cdot B^E &= & \boxed{\pm} \boxed{M} \boxed{M} \boxed{M} \cdot 10^{\boxed{\pm} \boxed{E} \boxed{E}} \\& & m = 3, \quad n = 2 \\M_{min} &= 1 &= & \pm 1.00 \\M_{max} &= B - B^{-(m-1)} &= & \pm 9.99 = 10 - 10^{-(3-1)} \\E_{min} &= -B^n + 1 &= & -99 = -10^2 + 1 \\E_{max} &= B^n - 1 &= & +99 = 10^2 - 1 \\x_{min} &= M_{min} \cdot B^{E_{min}} = B^{E_{min}} &= & 10^{-99} \\x_{max} &= M_{max} \cdot B^{E_{max}} &= & (10 - 10^{-(3-1)}) \cdot 10^{99} \\&= (B - B^{-(m-1)}) \cdot B^{E_{max}} < B^{E_{max}+1} &= & 9.9 \cdot 10^{99}\end{aligned}$$

Znaczące (istotne) i poprawne cyfry ułamkowe

Cyfry ułamkowe - wszystkie cyfry po kropce dziesiętnej

Cyfry istotne - wszystkie cyfry z wyjątkiem zer na początku liczby pomagające określić pozycję kropki

Przykład

123.12 – 2 cyfry ułamkowe, 5 cyfr istotnych

0.001234 – 6 cyfr ułamkowych, 4 cyfry istotne

Cyfry poprawne - cyfry ułamkowe dla których $|x - \tilde{x}| < \frac{1}{2} \cdot 10^{-t}$

Cyfry znaczące - cyfry istotne występujące do t -tej pozycji po kropce

Przykład

0.001234 ± 0.0000004 – 5 cyfr poprawnych, 3 cyfry znaczące ..., bo
 $0.0000004 = 0.4 \cdot 10^{-6} < \frac{1}{2} \cdot 10^{-6} \quad (x \in [0.001230, 0.001238])$

0.001234 ± 0.0000006 – 4 cyfry poprawne, 2 cyfry znaczące ..., bo
 $0.0000006 = 0.6 \cdot 10^{-6} < \frac{1}{2} \cdot 10^{-5} \quad (x \in [0.001240, 0.001228])$