

Metody numeryczne II. Wstęp. Teoria błędów.

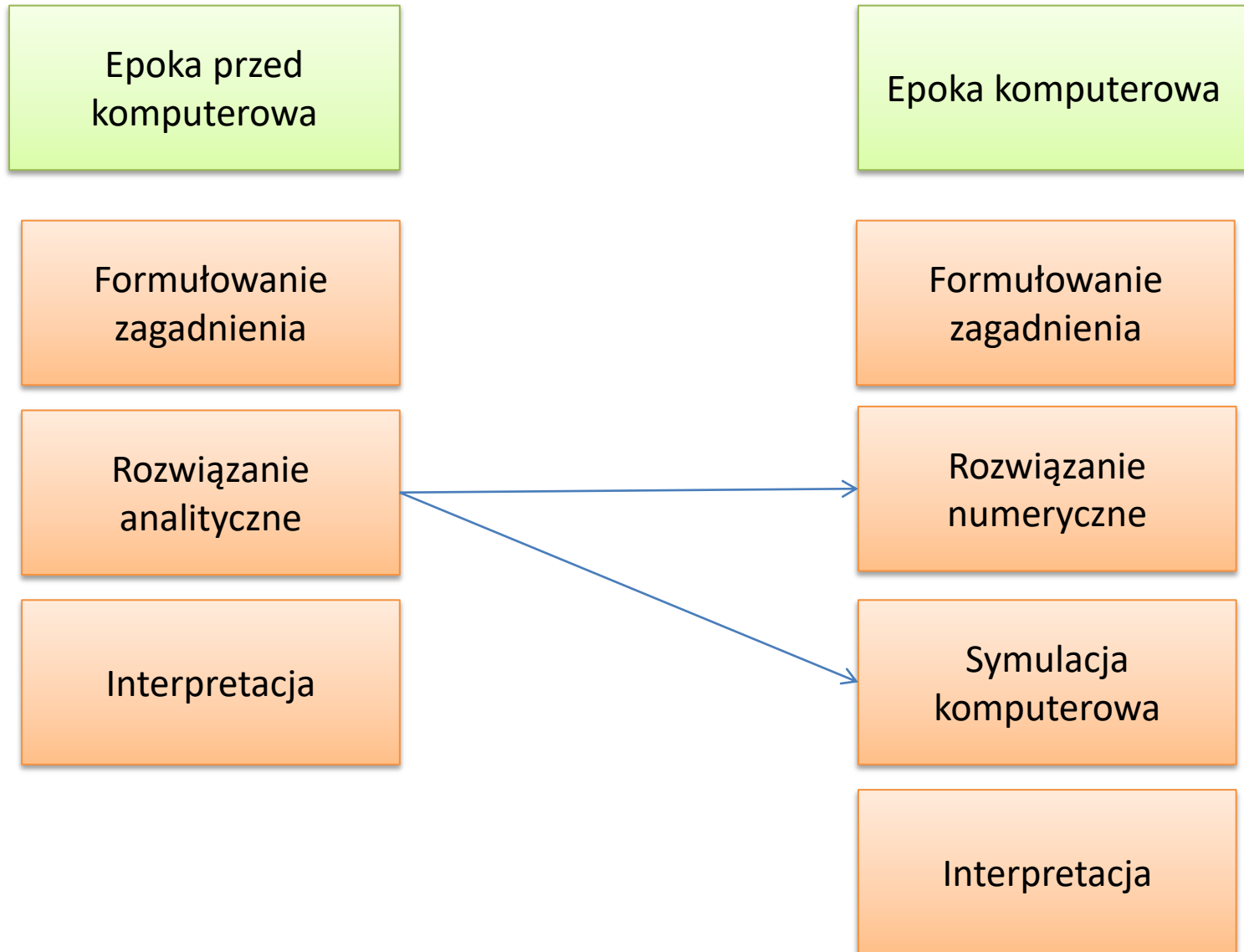
Oleksandr Sokolov

Wydział Fizyki, Astronomii i Informatyki Stosowanej UMK
<http://fizyka.umk.pl/~osokolov/MNII/>

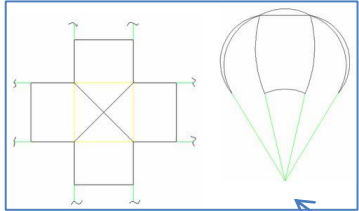
Dlaczego są metody numeryczne?

- Metody numeryczne są niezwykle potężnym narzędziem rozwiązywania problemów **modelowania natury**.
- Można korzystać z komercyjnych programów komputerowych, które mają **biblioteki metod numerycznych**. Trzeba znać podstawowe zasady metod numerycznych, aby być lepszym użytkownikiem takich programów.
- Niestety, są problemy, które nie mogą być rozwiązane poprzez istniejące programy komercyjne; wtedy trzeba **napisać swój własny program** odpowiedniej metody numerycznej.
- Metody numeryczne stanowią dobrą okazję do wzmocnienia **zrozumienia natury badanego zjawiska**.
- Te metody są potrzebne **w codziennej praktyce inżyniera lub naukowca**

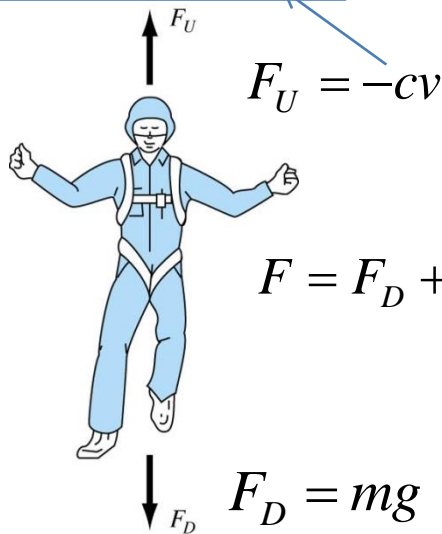
Uzasadnienie



Kiedy i na jakiej wysokości otwierać spadochron?



- F_D - **przyciąganie ziemskie** - powodujące stałe przyspieszenie wynoszące **9,81 m/s²**, a w konsekwencji ruch pionowy ku ziemi,
- F_U - **siła oporu powietrza** - skierowana przeciwnie do kierunku ruchu, zmniejszająca wpływ przyspieszenia na prędkość.



$$F = F_D + F_U$$

$$a = \frac{F}{m}$$

$$\frac{dv}{dt} = \frac{mg - cv}{m}$$

średnia utrata wysokości po określonym czasie:

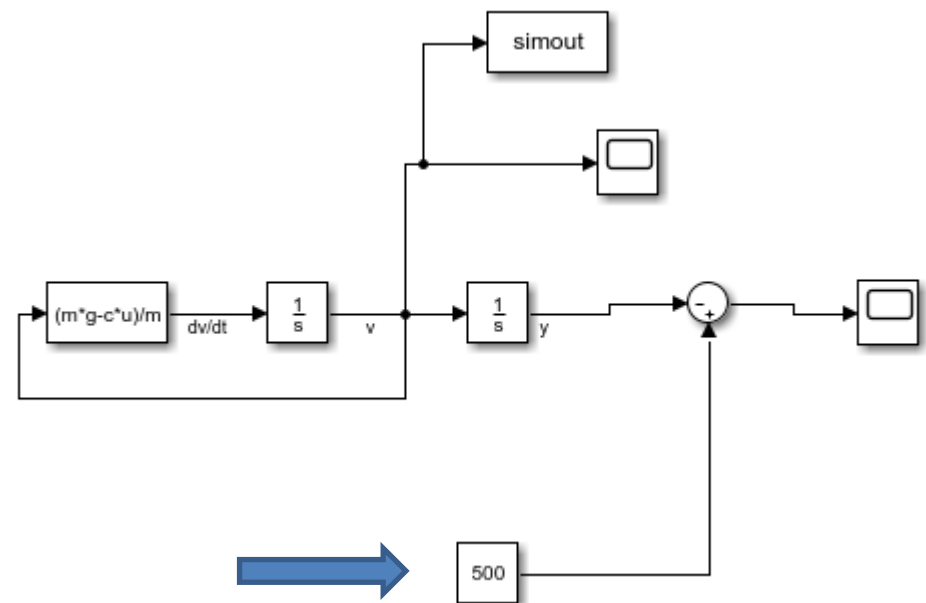
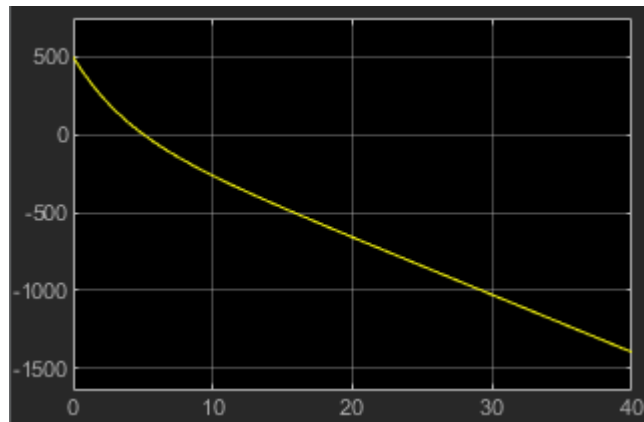
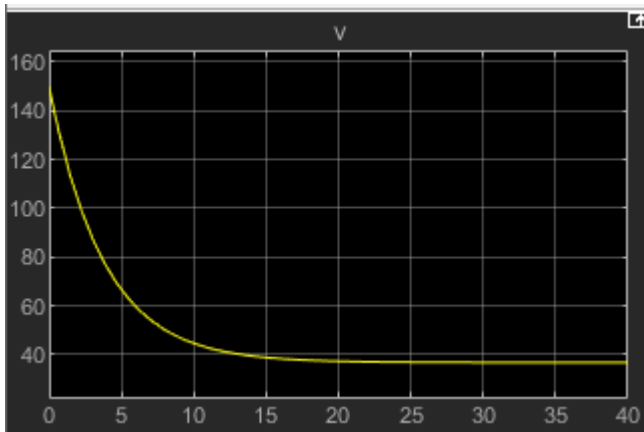
1 sek. - 5 m; 2 sek. - 20 m; 3 sek. - 45 m; 4 sek. - 75 m
5 sek. - 110 m; 6 sek. - 147 m; 7 sek. - 193 m; 8 sek. - 240 m
9 sek. - 288 m; 10 sek. - 340m

Wartość prędkości granicznej waha się w zależności od pozycji opadającego skoczka, ale do obliczeń przyjmuje się średnią prędkość, która wynosi 50 m/s (180 km/h).

Rozwiązanie analityczne

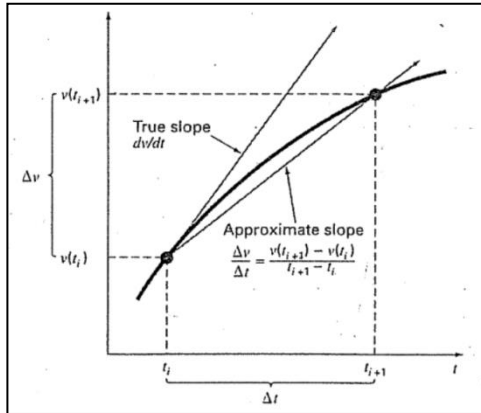
$$\frac{dv}{dt} = \frac{mg - cv}{m}$$

$$v(t) = \frac{mg}{c} (1 - e^{-(c/m)t})$$



\MNII\programOS\bledy\spadochron1.slx

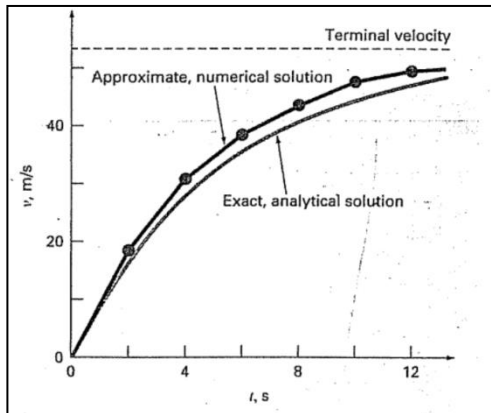
Rozwiązanie numeryczne



$$\frac{dv}{dt} = \frac{mg - cv}{m} = g - \frac{c}{m}v$$

$$\frac{dv}{dt} \cong \frac{\Delta v}{\Delta t} = \frac{v(t_{i+1}) - v(t_i)}{t_{i+1} - t_i}$$

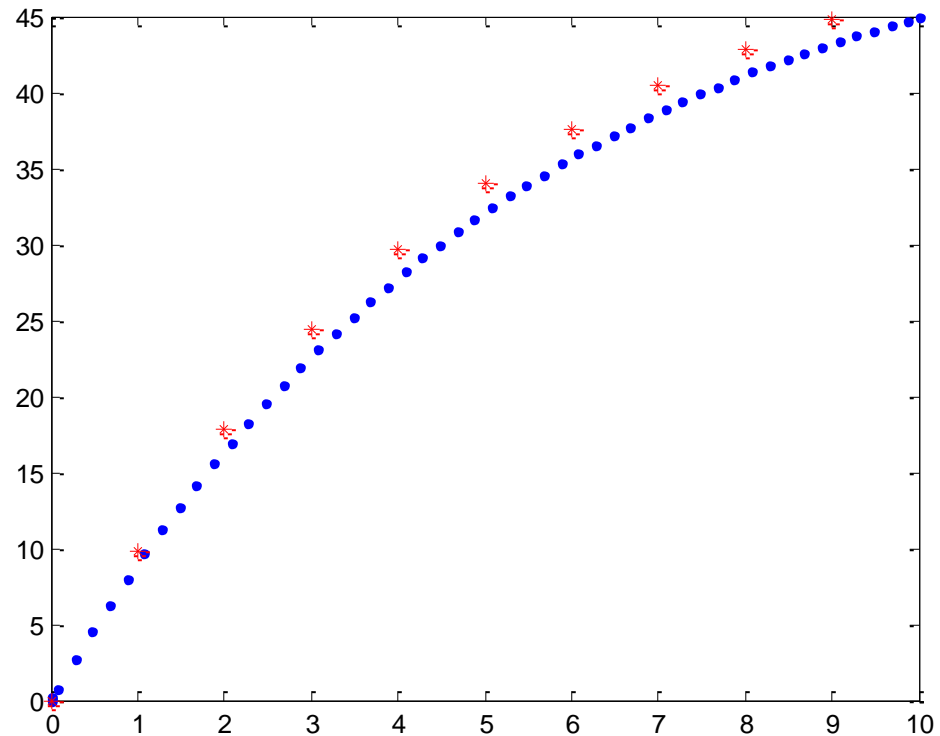
$$\frac{v(t_{i+1}) - v(t_i)}{t_{i+1} - t_i} = g - \frac{c}{m}v(t_i)$$



$$v(t_{i+1}) = v(t_i) + \left[g - \frac{c}{m}v(t_i) \right] (t_{i+1} - t_i)$$

Program obliczeń

```
vsim=[simout.Time simout.Data]
plot(vsim(:,1),vsim(:,2),'.')
g=9.81;%m/s^2
m=68.1;%kg
c=12.5;%kg/s
v=[];
t=0;
h=1;
hold on;
vnew=0;
for i=1:h:10;
    v=[v; (i-1)*h vnew];
    vnew=vnew+(g-c/m*vnew)*h;
end;
plot(v(:,1),v(:,2),'*r');
```



Problem:

Niska dokładność (duży krok) –
Duży czas obliczeń (mały krok)

Dynamika liczebności populacji

$$\frac{dN}{dt} = rN \left(1 - \frac{N}{K} \right),$$

gdzie:

N – liczebność populacji,

t – czas,

r – współczynnik tempa wzrostu

K – pojemność środowiska.



$$\frac{dx}{dt} = (a - by)x,$$

$$\frac{dy}{dt} = (cx - d)y,$$

gdzie:

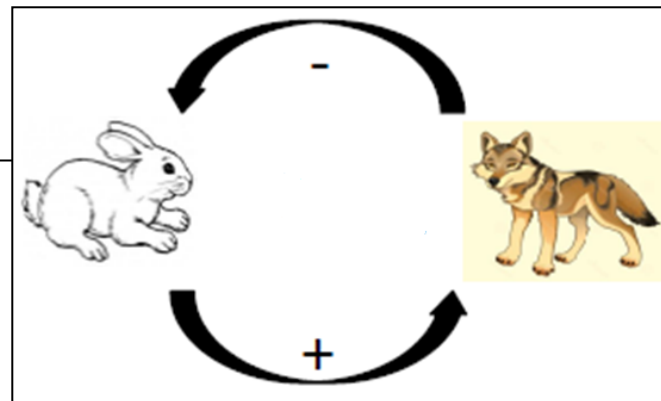
$x(t)$ – populacja, czyli liczba ofiar (ang. *prey*, np. *zająca*),

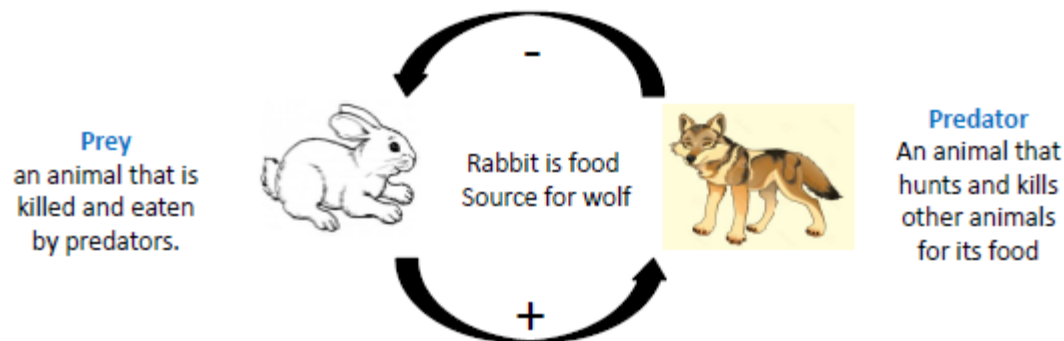
$y(t)$ – liczba drapieżców (ang. *predators*, np. *rysie*),

t – rozwój tych dwóch populacji w czasie,

stałe ($a, b, c, d > 0$, dodatnie parametry) oznaczają:

- a – częstość narodzin ofiar lub współczynnik przyrostu ofiar,
- b – częstość umierania ofiar na skutek drapieżnictwa,
- c – częstość narodzin drapieżników lub współczynnik przyrostu drapieżników,
- d – częstość umierania drapieżników lub współczynnik ubywania drapieżników,





Hypothesis

1. Assume that the system can support an unlimited amount of rabbit, so that in the absence of predation, the prey population grows exponentially.
2. In the absence of predation, the predator population dies off exponentially (no food supply).
3. Then the growth rate of the rabbit is diminished in a way that is proportional to the number of interactions between them and the wolf

$$x' = ax - \underbrace{bxy}_{\text{Likelihood of species meeting}} = (a - by)x$$

Likelihood of
species meeting

$$y' = -cy + \underbrace{dxy}_{\text{Likelihood of species meeting}} = (-c + dx)y$$

- If $y = 0$, $x' = ax$, \Rightarrow exponential growth
- If $x = 0$, $y' = -cy$, \Rightarrow exponential growth
- Steady-state: $(x, y) = (0, 0)$
- If $x \neq 0$, $a - by = 0$, $\Rightarrow y = a/b \neq 0$
- If $y \neq 0$, $-c + dx = 0$, $\Rightarrow x = c/d \neq 0$
- Steady-state: $(x, y) = (c/d, a/b)$

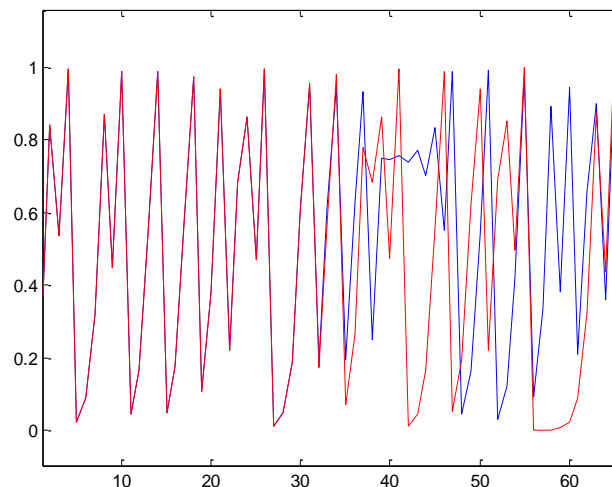
Chaos deterministyczny

Odwzorowanie logistyczne – funkcja odwzorowująca przedział jednostkowy w siebie: $f : [0, 1] \rightarrow [0, 1]$ dana przepisem:

$$x \mapsto kx(1 - x)$$
 przy warunku na wartości parametru $0 < k < 4$.

```
x=0.3;
out=[];
for i=1:300
    out=[out;i,x];
    x=2.4*x*(1-x);
end;
x=0.300000000001;
out1=[];
for i=1:300
    out1=[out1;i,x];
    x=2.4*x*(1-x);
end;
plot(out(:,1),out(:,2));
hold on;
plot(out1(:,1),out1(:,2),'r');
```

Chaos deterministyczny – własność równań lub układów równań, polegająca na dużej wrażliwości rozwiązań na dowolnie małe zaburzenie parametrów. Dotyczy to zwykle nieliniowych równań różniczkowych i różnicowych, opisujących układy dynamiczne.



Reprezentacja liczb

- W języku C wyróżniamy następujące typy zmiennych (wg wielkości) :
- 4 podstawowe :
 - **char** - jednobajtowe liczby całkowite, służy do przechowywania znaków;
 - **int**- typ całkowity, o długości domyślnej dla danej architektury komputera;
 - **float** - typ **zmiennopozycyjny** (zwany również zmiennoprzecinkowym), reprezentujący liczby rzeczywiste (4 bajty); Są dokładnie opisane w [IEEE 754](#).
 - **double** - typ **zmiennopozycyjny** podwójnej precyzji (8 bajtów);

FLOAT $3.4 * (10^{-38})$ to $3.4 * (10^{+38})$

Przykład

[*] w1OS1.c

```
1  #include <stdio.h>
2  #include <stdlib.h>
3
4      int main()
5  {
6      float a,b,c,d,e,f,g;
7      a=123456789;
8      b=a+1;
9      c=a+2;
10     d=a+3;
11     e=a+4;
12     f=a+5;
13     g=a+6;
14
15
16
17     printf("Result: a %f\n", a);
18     printf("Result: b %f\n", b);
19     printf("Result: c %f\n", c);
20     printf("Result: d %f\n", d);
21     printf("Result: e %f\n", e);
22     printf("Result: f %f\n", f);
23     printf("Result: g %f\n", g);
24
25     return 0;
26
27 }
```

The float type can represent values ranging from approximately 1.5×10^{-45} to 3.4×10^{38} , with a precision — the limit of digits — of **seven**. Float can contain up to seven digits in *total*, not just following the decimal point — so, for example, 321.1234567 cannot be stored in float because it has 10 digits. If greater precision—more digits—is necessary, the double type is used.

w1OS2.c

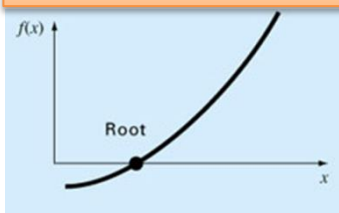
Przykład (cd)

```
a=123456789; ➡  
b=a+1;  
c=a+2;  
d=a+3;  
e=a+4;  
f=a+5;  
g=a+6;
```

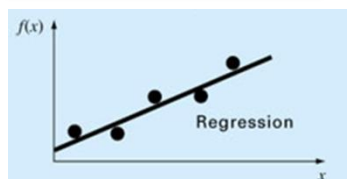
```
C:\Users\KIS-OS\OneDrive\saszasokolov\Nauka  
Result: a 123456792.000000  
Result: b 123456792.000000  
Result: c 123456792.000000  
Result: d 123456792.000000  
Result: e 123456800.000000  
Result: f 123456800.000000  
Result: g 123456800.000000  
-----  
Process exited with return value 0  
Press any key to continue . . .
```

Podstawowe problemy obliczeniowe

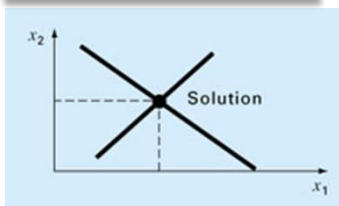
Korzeń równania



Aproksymacja



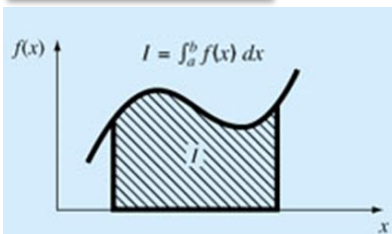
Układ równań



Interpolacja



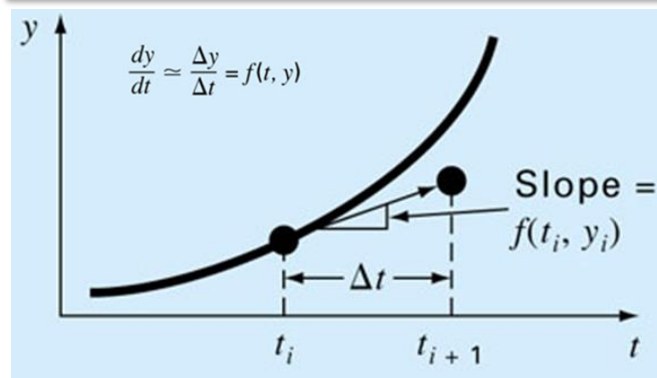
Całkowanie



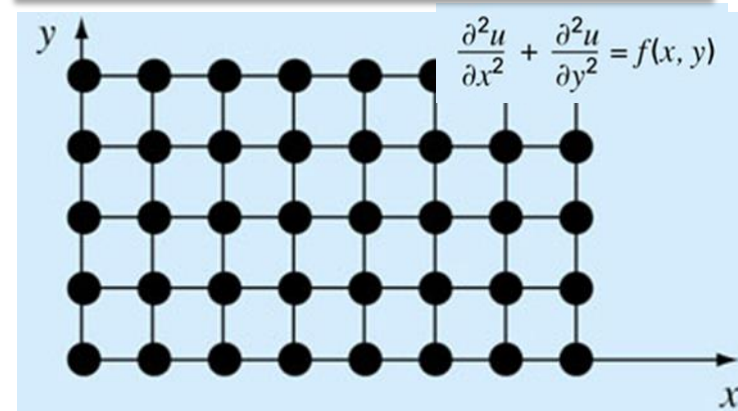
Optymalizacja



Równania różniczkowe zwyczajne



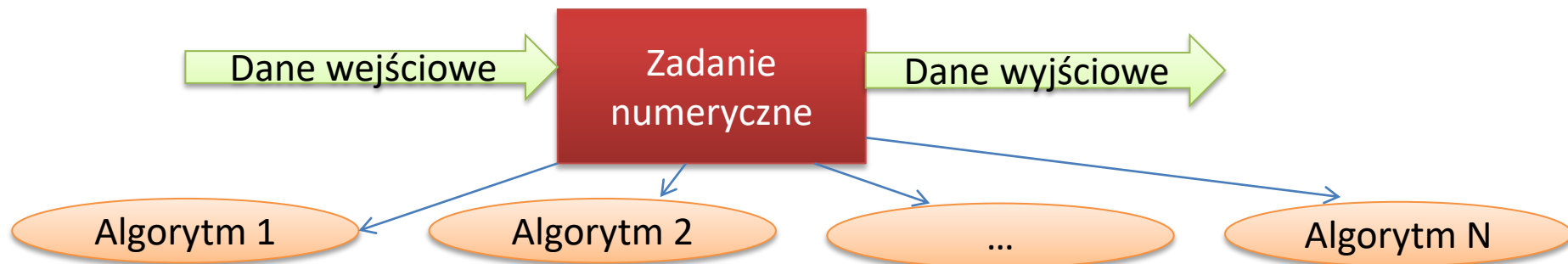
Równania różniczkowe cząstkowe



Metody numeryczne a błędy

Przez **zadanie numeryczne** rozumiemy jasny i jednoznaczny opis powiązania funkcjonalnego między **danymi wejściowymi** (zmiennymi niezależnymi) i **danymi wyjściowymi** (szukanymi wynikami). **Algorytm** dla danego zadania numerycznego jest z definicji pełnym opisem poprawnie określonych **operacji przekształcających dopuszczalne dane wejściowe na dane wyjściowe**.

„Operacje” oznaczają tu **działania arytmetyczne i logiczne**. Dla danego zadania numerycznego można rozważać wiele różnych algorytmów. Algorytmy te mogą dawać wyniki o **bardzo różnej dokładności** (w AiSD – złożoność)



Błędy obliczeń

W teorii metod numerycznych zasadniczą rolę odgrywa zrozumienie **ograniczeń danej metody**, co jest ściśle związane z określeniem **błędu obliczeniowego**.

Błędy modelu - **uproszczenie modelu matematycznego** (przyjęcie założeń upraszczających),

Błędy danych wejściowych (gdy wykorzystujemy dane zaokrąglone, pochodzące np. z wcześniejszych obliczeń lub gdy dane wejściowe są wynikiem pomiarów wielkości fizycznych obarczonych **błędami pomiarowymi**),

Błędy aproksymacji – np. **pochoďnej** za pomocą ilorazu różnicowego, obliczanie wartości **całki** oznaczonej jako granicy sum przybliżających ją podziałów itp.).

Błędy zaokrągleń (związane z odpowiednią **reprezentacją liczby**),

Błędy obcięcia (gdy proces obliczania granicy jest przerywany przed **osiągnięciem wartości granicznej** - np. ograniczenie szeregu nieskończonego do skończonej liczby składników,

Miary błędów

- Błąd bezwzględny

$$\Delta = a - \tilde{a}$$

Błąd bezwzględny informuje, o ile różni się wartość zmierzona od dokładnej

- Błąd względny

$$\delta = \frac{\Delta}{a} \cdot 100\%$$

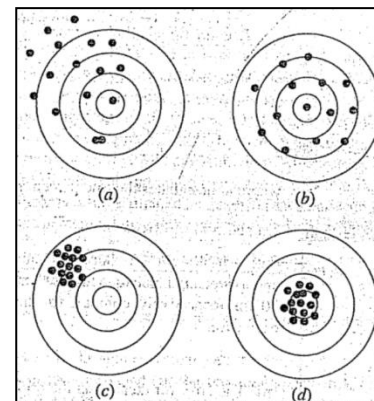
Czasami błąd względny wyrażamy w procentach



dokładna wartość -
Przybliżona
(zmierzona) wartość -

a
 \tilde{a}

Nieokreśloność a niedokładność



Przykład



$$a = 10000m$$

$$\tilde{a} = 9999m$$

$$\Delta = a - \tilde{a} = 1m$$

$$\delta = 1 / 10000 * 100\% = 0.01\%$$



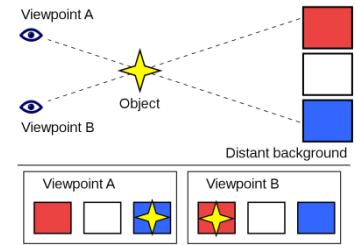
$$a = 5cm$$

$$\tilde{a} = 4cm$$

$$\Delta = a - \tilde{a} = 1cm$$

$$\delta = 1 / 5 * 100\% = 20\%$$

Błędy pomiarów



Każde doświadczenie fizyczne wymaga przeprowadzenia **oszacowania błędu**, którym jest obarczony wynik tzn. podania z jaką dokładnością dana wielkość została wyznaczona.

Ogólnie rozróżniamy:

- **błędy grube** wynikłe z nieuwagi i z pomyłek eksperymentatora (np. przy odczycie lub w zapisie wyniku). Często są jednorazowe i bardzo duże.
- **błędy systematyczne** wynikłe ze złego (mało dokładnego) ustawienia samego eksperymentu
 - **nie uwzględnienie pewnych poprawek** np. siły wyporu powietrza przy dokładnym ważeniu),
 - **wad urządzeń pomiarowych** (przykładem może być waga dźwigniowa z przesuniętym punktem zawieszenia, czasomierz wskazówkowy ze środkiem skali nie pokrywającym się z osią wskazówek czy źle wyskalowane przyrządy),
 - **ze stanu zewnętrznych warunków pomiaru** (np. zbyt wysoka temperatura w pomieszczeniu) jak i z błędu eksperymentatora (np. znany *błąd paralaksy*).
- **błędy przypadkowe** wynikłe z niedokładności odczytu, fluktuacji warunków pomiaru, z nieokreślenia samej mierzonej wielkości fizycznej itp.

Nie można wykonać bezbłędnego wyznaczenia wielkości fizycznej tzw. pomiaru absolutnie dokładnego

Probabilistyczna teoria błędów Gaussa

Z jednego pomiaru nie możemy wnioskować o jego dokładności. Do tego konieczna jest ich seria. **Wartości rzeczywistej nie znamy.**

Z serii pomiarów wartością **najbardziej zbliżoną do wartości rzeczywistej** jest średnia arytmetyczna

$$\bar{X} = \frac{1}{N} \sum_{i=1}^N X_i$$

Jest to podstawowe twierdzenie teorii błędów tzw. **pierwszy postulat Gaussa**.

Wynika ono z **faktu równości prawdopodobieństw** jak zawyżenia wielkości mierzonej tak i jej zaniżenia.

Tym samym wartość średnia \bar{X} jest **jedynie blisko położona wartości rzeczywistej X_R** , ale nie równa jej.

Równość $\bar{X} = X_R$ moglibyśmy napisać tylko **dla serii nieskończenie długiej** pomiarów

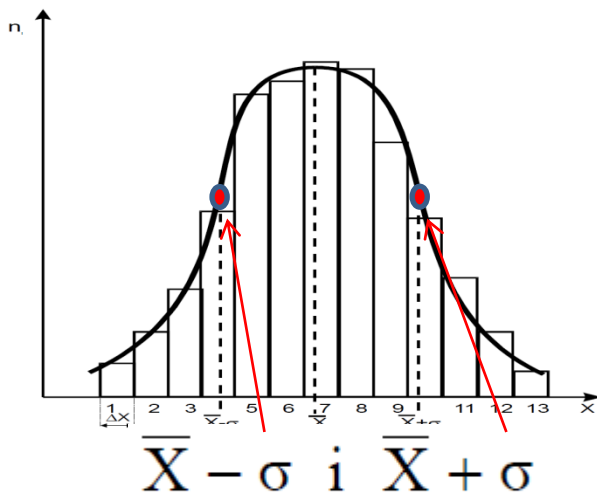
Rozkład Gaussa

Wyniki pomiarów w serii rozkładają się **wokół wartości średniej** w tzw. krzywą Gaussa

$$P(X) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(X - \bar{X})^2}{2\sigma^2}}$$

Odchylenie standardowe określa rozmycie rozkładu wokół **wartości średniej**

Aby się o tym przekonać należy zakres pomiarowy podzielić na **przedziały o równej szerokości ΔX** i obliczyć **ile pomiarów** z serii zmieściło się w każdym z nich



Zauważmy, że na krzywej Gaussa można wyróżnić obszary o **przeciwnie skierowanej krzywiznie**. W okolicy maksimum krzywa jest **wypukła**, a daleko poza maksimum staje się krzywą **wklęstą**.

Oczywiście obszary takie są oddzielone **punktami przegięcia**.

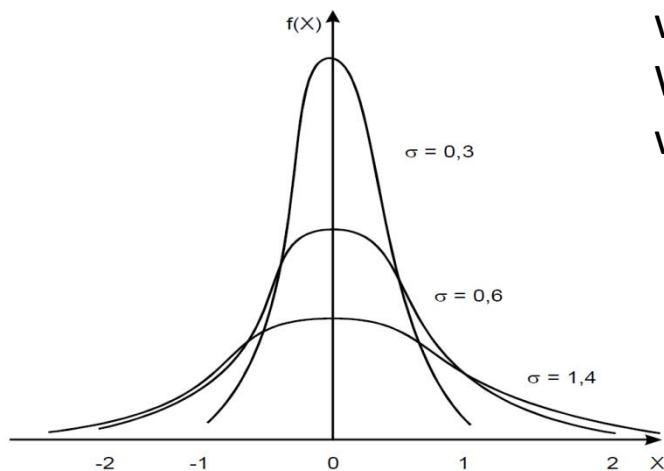
Przedział $\langle X - \sigma, X + \sigma \rangle$

Ponieważ rozkład Gaussa opisuje **zjawisko probabilistyczne**, a więc można określić jedynie **prawdopodobieństwo**, że dowolny **wynik pomiaru** X_i ($i=1,2,3,\dots,N$) znajdzie się w aktualnie interesującym nas przedziale wartości $\langle X_a, X_b \rangle$.

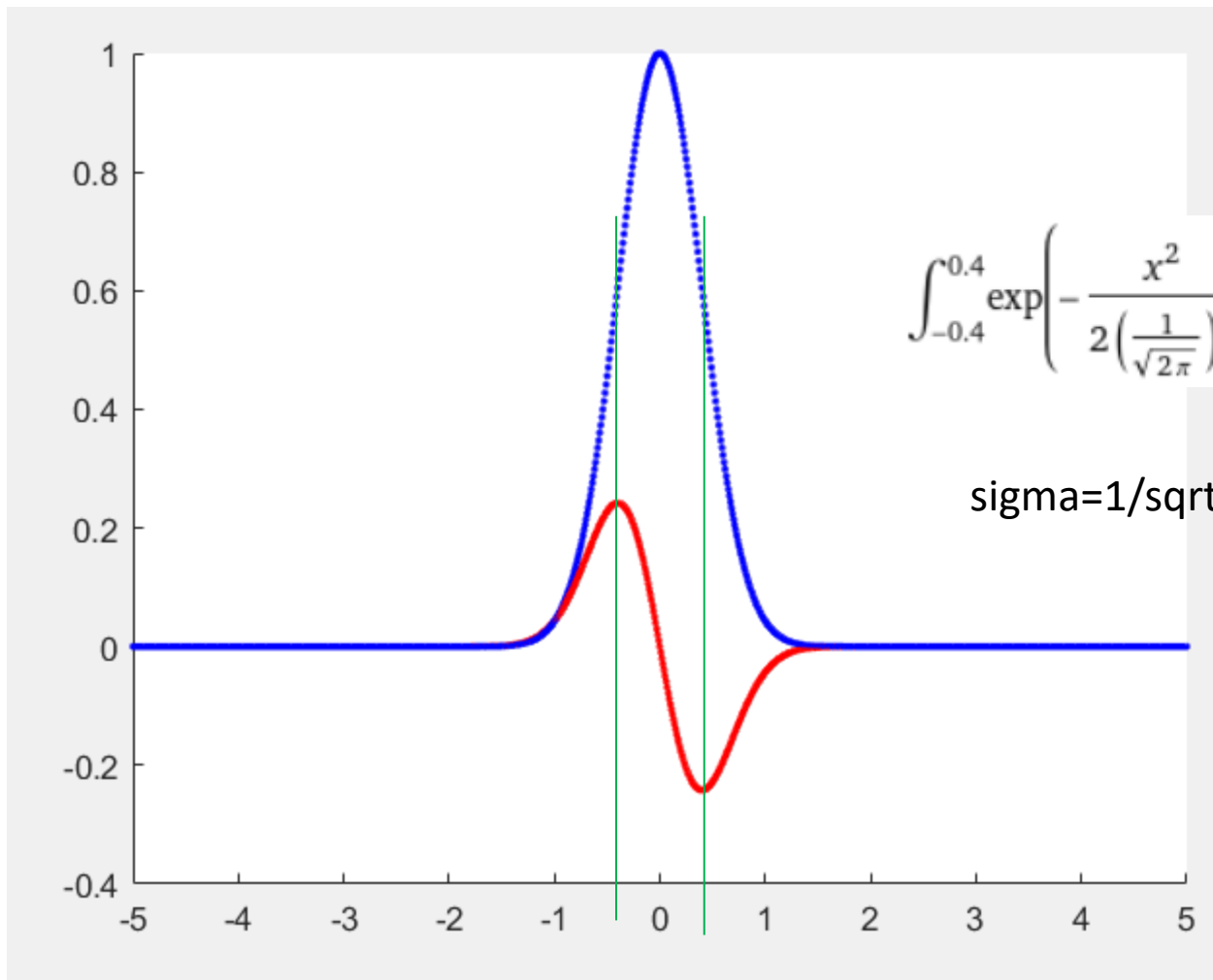
W przedziale $\langle X - \sigma, X + \sigma \rangle$ mieści się 68,26% wyników z serii

W przedziale $X - 2\sigma, X + 2\sigma$ mieści się 95,45% wyników z serii.

W przedziale $X - 3\sigma, X + 3\sigma$ mieści się 99,73% wyników z serii.



Dla tego im większe jest odchylenie standardowe, tym szersza jest krzywa i bardziej spłaszczona



$$\int_{-0.4}^{0.4} \exp\left(-\frac{x^2}{2\left(\frac{1}{\sqrt{2\pi}}\right)^2}\right) dx = \underline{0.683971}$$

$$\text{sigma} = 1/\sqrt{2 \cdot \pi} = 0.3989;$$

Średni błąd kwadratowy

Odchylenie standardowe w teorii błędów nazywa się **średnim błędem kwadratowym** i oblicza się go z wyrażenia:

$$\sigma = \sqrt{\frac{\sum_{i=1}^N (X_i - \bar{X})^2}{(N-1)}}.$$

Występujący w tym wyrażeniu czynnik $N - 1$ można uzasadnić w ten sposób, że ponieważ część informacji zawartej w serii je $X_1, X_2, X_3, \dots, X_N$ została wykorzystana do określenia wartości średniej \bar{X} , uśrednianie związane z odchyleniem standardowym następuje z mniejszą liczbą punktów swobody i stąd podzielenie przez $N - 1$ zamiast przez N .

Najczęściej wyznaczany jest jednak jako optymalny średni błąd kwadratowy σ , a z niego **średni błąd kwadratowy wartości średniej**:

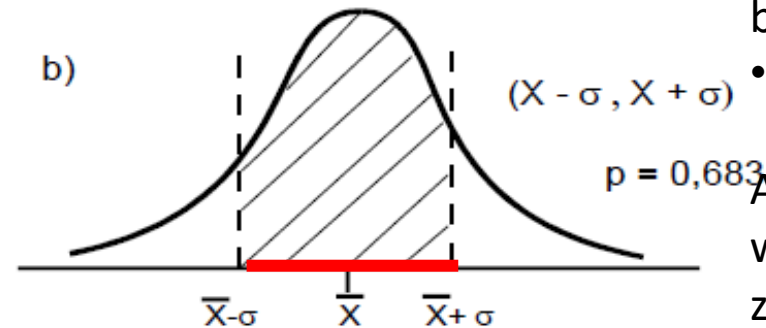
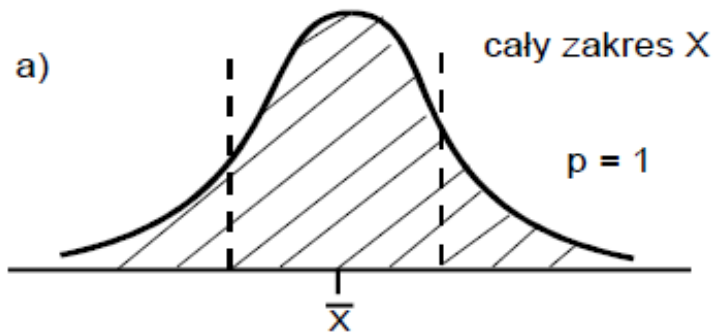
$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{N}} \quad \longrightarrow \quad \sigma_{\bar{X}} = \sqrt{\frac{\sum_{i=1}^N (X_i - \bar{X})^2}{N(N-1)}}$$

Błąd średni kwadratowy jest najważniejszym i najczęściej stosowanym **wskaźnikiem dokładności pomiaru**.

Przedziały ufności i poziomy ufności

Prawdopodobieństwo, że dany pomiar z serii pomiarowej znajdzie się w przedziale $X - \sigma$, $X + \sigma$ wynosi **0,683**.

W interpretacji graficznej prawdopodobieństwu temu odpowiada pole pod krzywą Gaussa odcięte tym przedziałem przy założeniu, że **pole pod całą krzywą równa się jeden** (rys. ab).



poziom ufności – 0.683

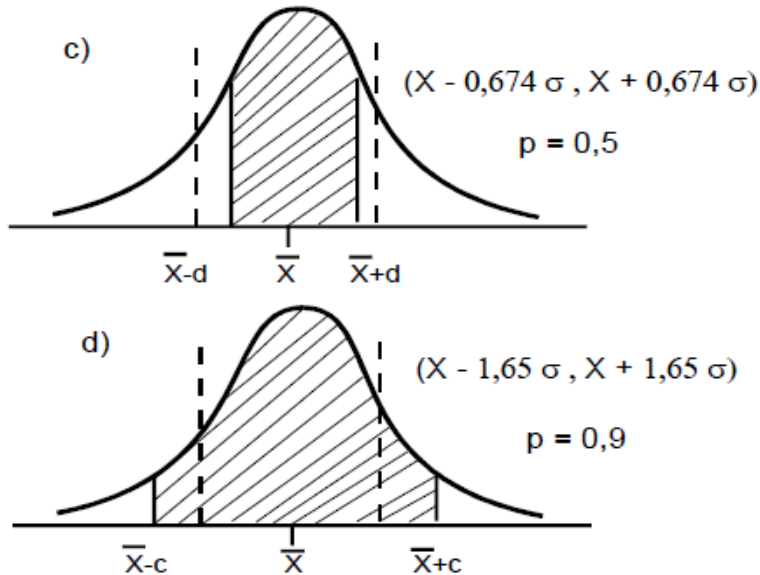
przedział ufności

W eksperymencie oczywiście chcielibyśmy, żeby

- błąd wyniku (**przedział ufności**) był jak najmniejszy przy możliwie dużym
- prawdopodobieństwie (**poziomie ufności**).

Analizując kształt krzywej dzwonowej dochodzimy do wniosku, że optymalność przedziału $X - \sigma$, $X + \sigma$ wynika z faktu, że **jest on wyznaczony przez punkty przegięcia krzywej**.

Przedziały ufności i poziomy ufności (cd)



Gdybyśmy chcieli sztucznie zmniejszyć ten **przedział ufności** do $\bar{X} - d$, $\bar{X} + d$ (rys. c), to znacznie **stracimy na poziomie ufności** (o pole pod krzywą Gaussa odcięte przedziałami $\bar{X} - \sigma$, $\bar{X} - d$, $\bar{X} + d$, $\bar{X} + \sigma$, które jest duże, bo na tych odcinkach krzywa dzwonowa jest wypukła).

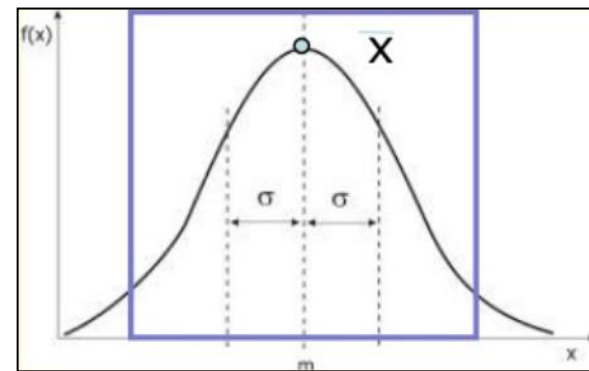
Gdybyśmy z kolei chcieli sztucznie **podnieść poziom ufności** (rys. d), to jest to możliwe tylko przez **znaczne poszerzenie** tego przedziału ufności do $\bar{X} - c$, $\bar{X} + c$, gdyż pola pod krzywą w przedziałach oddalonych od średniej \bar{X} dalej niż o σ **wnoszą już mały wkład do poziomu ufności** (krzywa jest tu wklęsła).

Poziom (współczynnik) ufności

$$P\left\{\bar{X} - t_{\alpha} \frac{s}{\sqrt{n}} < m < \bar{X} + t_{\alpha} \frac{s}{\sqrt{n}}\right\} = 1 - \alpha$$

n pomiarów

s - odchylenie standardowe z próby
 $t_{\alpha, n-1}$ - wartość odczytana z tablic rozkładu t-Studenta
 α - współczynnik ufności, 0-1
 m - wartość zmierzona



$$X_1, \dots, X_n \quad N(\mu, \sigma^2) \quad N(0,1)$$

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$



$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

$$\frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

$$\frac{\bar{X} - \mu}{s / \sqrt{n}}$$

Jest rozkładem Studenta

1. **Poziom ufności** – mówi o tym, jak bardzo możemy „zaufać” naszym wynikom, więc jest to duża wartość (uwzględniając fakt, że jest z przedziału (0,1), więc przez słowo duża rozumiemy bliską do wartości 1). Jest to $1 - \alpha$.
2. **Poziom istotności** – mówi nam o tym na jak dużą pomyłkę sobie pozwalamy (na ile procent możemy się mylić). Jest to α .

Tabela rozkładu T



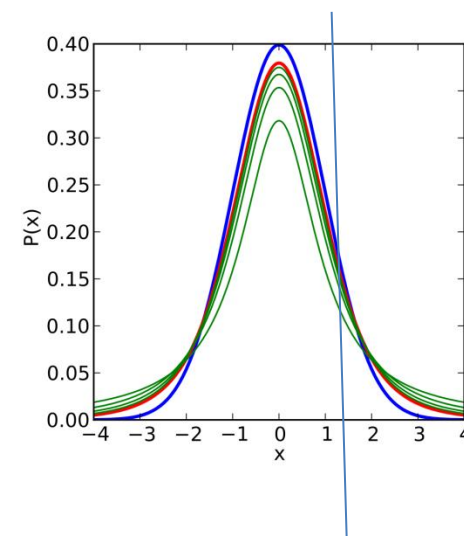
| r | 75 % | 80 % | 85 % | 90 % | 95 % | 97.5 % | 99 % | 99.5 % | 99.75 % | 99.9 % | 99.95 % |
|-----|-------|-------|-------|-------|-------|--------|-------|--------|---------|--------|---------|
| 1 | 1.000 | 1.376 | 1.963 | 3.078 | 6.314 | 12.71 | 31.82 | 63.66 | 127.3 | 318.3 | 636.6 |
| 2 | 0.816 | 1.061 | 1.386 | 1.886 | 2.920 | 4.303 | 6.965 | 9.925 | 14.09 | 22.33 | 31.60 |
| 3 | 0.765 | 0.978 | 1.250 | 1.638 | 2.353 | 3.182 | 4.541 | 5.841 | 7.453 | 10.21 | 12.92 |
| 4 | 0.741 | 0.941 | 1.190 | 1.533 | 2.132 | 2.776 | 3.747 | 4.604 | 5.598 | 7.173 | 8.610 |
| 5 | 0.727 | 0.920 | 1.156 | 1.476 | 2.015 | 2.571 | 3.365 | 4.032 | 4.773 | 5.893 | 6.869 |
| 6 | 0.718 | 0.906 | 1.134 | 1.440 | 1.943 | 2.447 | 3.143 | 3.707 | 4.317 | 5.208 | 5.959 |
| 7 | 0.711 | 0.896 | 1.119 | 1.415 | 1.895 | 2.365 | 2.998 | 3.499 | 4.029 | 4.785 | 5.408 |

$n=4$, 90%

$$P(T < 1.533) = 0.9 \text{ (całka rozkładu)}$$

$$P(T < -1.533) = P(T > 1.533) = 1 - 0.9 = 0.1$$

$$P(-1.533 < T < 1.533) = 1 - 2 * 0.1 = 0.8$$



Przykład

x=[2 3 4 4 3 3 3];

Xs=3.6666 sigma=0.8924

Współczynnik ufności:

$$P = 0,975 = 1 - 0,05 \Rightarrow \alpha = 0,05$$

$$P \left\{ \bar{X} - t_{\alpha} \frac{s}{\sqrt{n}} < m < \bar{X} + t_{\alpha} \frac{s}{\sqrt{n}} \right\} = 1 - \alpha$$

$$3.6666 - 2.365 * 0.8924 / \sqrt{7} = 2.8689$$

$$3.6666 + 2.365 * 0.8924 / \sqrt{7} = 4.4643$$

$$P(2.8689, 4.4643) = 0.975$$

Błędy danych

Cyfry ułamkowe to wszystkie cyfry *po kropce dziesiętnej w ułamku*.

Cyfry istotne to cyfry jakie pozostaną *po pominięciu zer na początku*.

Przykład

Liczba 0,00245 ma pięć cyfr ułamkowych i trzy istotne.

Liczba 12,12 ma dwie cyfry ułamkowe i cztery istotne.

Cyfry znaczące – cyfry rozwinięcia dziesiętnego mierzonej wielkości fizycznej, począwszy *od pierwszej cyfry niezerowej aż do ostatniej cyfry*, której wartość nie zmienia się wewnątrz przyjętego **przedziału ufności** ε .

Przyjmijmy, że zapis $x = \tilde{x} \pm \varepsilon$ oznacza, że $|\tilde{x} - x| \leq \varepsilon$. Wartość $\varepsilon = \max|x - \tilde{x}|$ nazywamy *maksymalnym błędem bezwzględnym* lub *błędem granicznym*.

Jeśli $|\tilde{x} - x| \leq \frac{1}{2} \cdot 10^{-t}$, to mówimy, że \tilde{x} ma t *poprawnych cyfr ułamkowych*. Cyfry istotne występujące aż do pozycji t -tej po kropce nazywamy *cyframi znaczącymi*.

BŁĘDY POMIAROWE A ZAOKRĄGLANIE WYNIKÓW

W ogólnym przypadku wynik pomiaru przedstawiany jest w postaci:

$$X_R = X_M \pm \Delta X$$

Gdzie:

X_R - wartość rzeczywista wielkości mierzonej

X_M - wartość uzyskana w wyniku pomiaru

ΔX - niepewność lub błąd pomiaru

Powyższy zapis oznacza, że:

- najlepszym przybliżeniem wartości mierzonej jest wg eksperymentatora liczba X_M

- **z rozsądnym prawdopodobieństwem** szukana wartość znajduje się pomiędzy

$$X_M - \Delta X \text{ i } X_M + \Delta X$$

Błąd pomiaru ΔX jest wartością oszacowaną.

Obliczone wartości X_M i ΔX podajemy zaokrąglone tzn. przybliżamy wartości otrzymane z obliczeń zgodnie z powszechnie przyjętymi zasadami zaokrągleń.

Cyfry pewne. Jeżeli błąd spowodowany przybliżeniem liczby dziesiętnej jest mniejszy od jedności na ostatnim miejscu dziesiętnym to mówimy, że wszystkie jej cyfry są pewne. Przybliżenie dziesiętne podaje się wtedy z zachowaniem tylko cyfr pewnych np. $125 \cdot 10^3$

lub $1,25 \cdot 10^5$.

Zaokrąglanie liczb

Jeżeli cyfra na t+1-szej pozycji jest mniejsza niż 5 to wynik bez zmian, gdy jest większa od 5 to t-tą cyfrę zwiększamy o 1. Natomiast w przypadku gdy t+1-sza cyfra równa się 5 to gdy t-ta cyfra jest parzysta to bez zmian a jeżeli nieparzysta to zwiększamy o 1.

0.312

Jeżeli cyfra 2 jest poprawna, to błąd bezwzględny musi wynosić +/-0.0005

Czyli

$0.312 \pm 0.0005 = \{0.3115 \ 0.3116 \ 0.3117 \ \dots \ 0.3125\}$

Umieszczenie cyfry na najmniej znaczącym miejscu oznacza, że błąd bezwzględny wyznaczonej liczby jest **mniejszy lub równy połowie rzędu tego miejsca**.

0.34567 zaokrąglamy do 0.3457

0.34565 zaokrąglamy do 0.3456

0.34575 zaokrąglamy do 0.3458

Przykład

$$x = \tilde{x} \pm \varepsilon$$

$$\varepsilon = \max|x - \tilde{x}|$$

$$|\tilde{x} - x| \leq \frac{1}{2} \cdot 10^{-t}$$

0.00147 – 5 cyfr ułamkowych, 3 cyfry istotne,

12.34 – 2 cyfry ułamkowe, 4 cyfry istotne,

0.001234 ± 0.000004 – 5 cyfr poprawnych, 3 cyfry znaczące,

0.001234 ± 0.000006 – 4 cyfry poprawne, 2 cyfry znaczące.

$$0.000004 < 1/2 * 10^{-5}$$

$$>> 0.000004 < 1/2 * 10^{-5} \quad \mathbf{1}$$

$$>> 0.000006 < 1/2 * 10^{-5} \quad \mathbf{0}$$

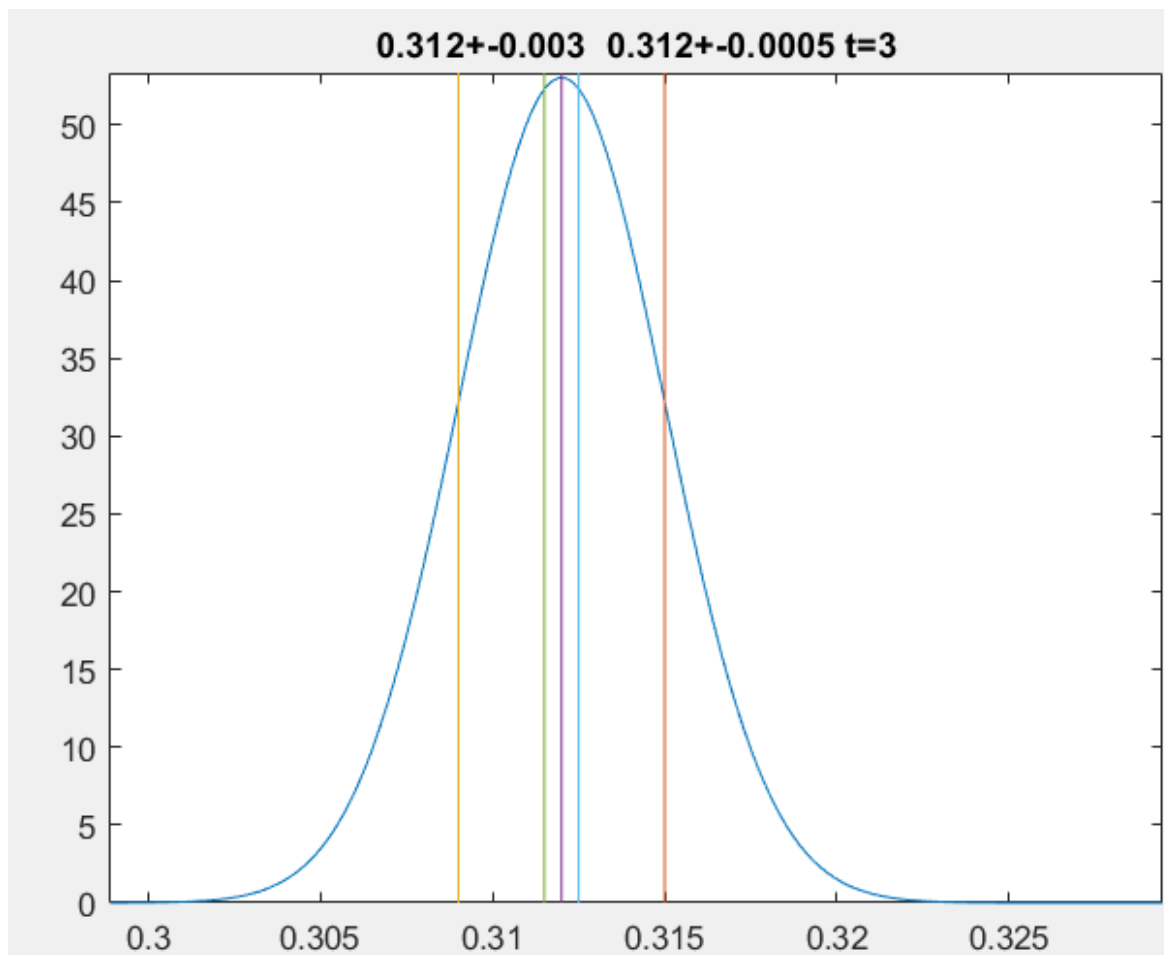
$$>> 0.000006 < 1/2 * 10^{-4} \quad \mathbf{1}$$

Cyfry znaczące – cyfry rozwinięcia dziesiętnego mierzonej wielkości fizycznej, począwszy **od pierwszej cyfry niezerowej aż do ostatniej cyfry**, której wartość nie zmienia się wewnątrz przyjętego **przedziału ufności** ε .

Cyfra 2 jest poprawna

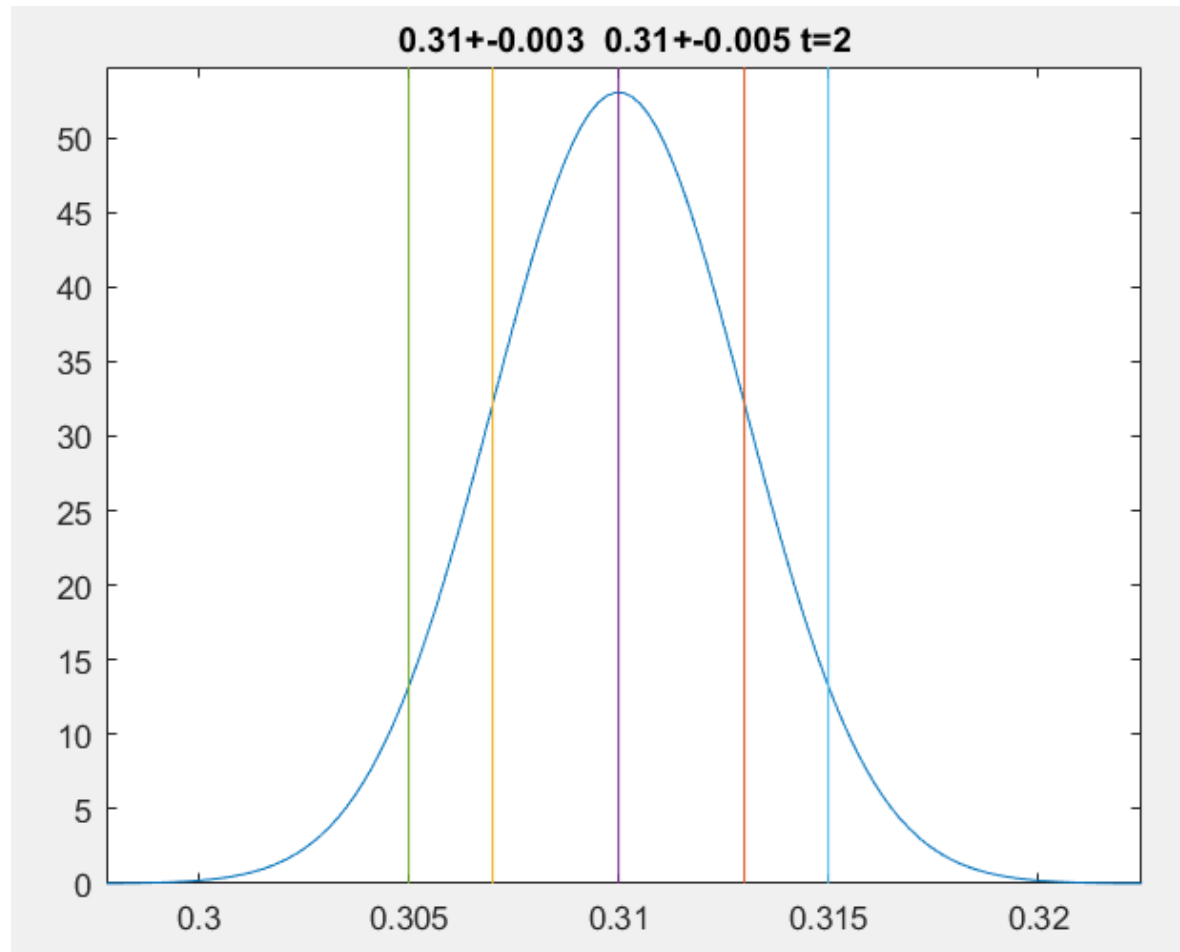
sprzeczność

0.31²



Cyfra 1 jest poprawna

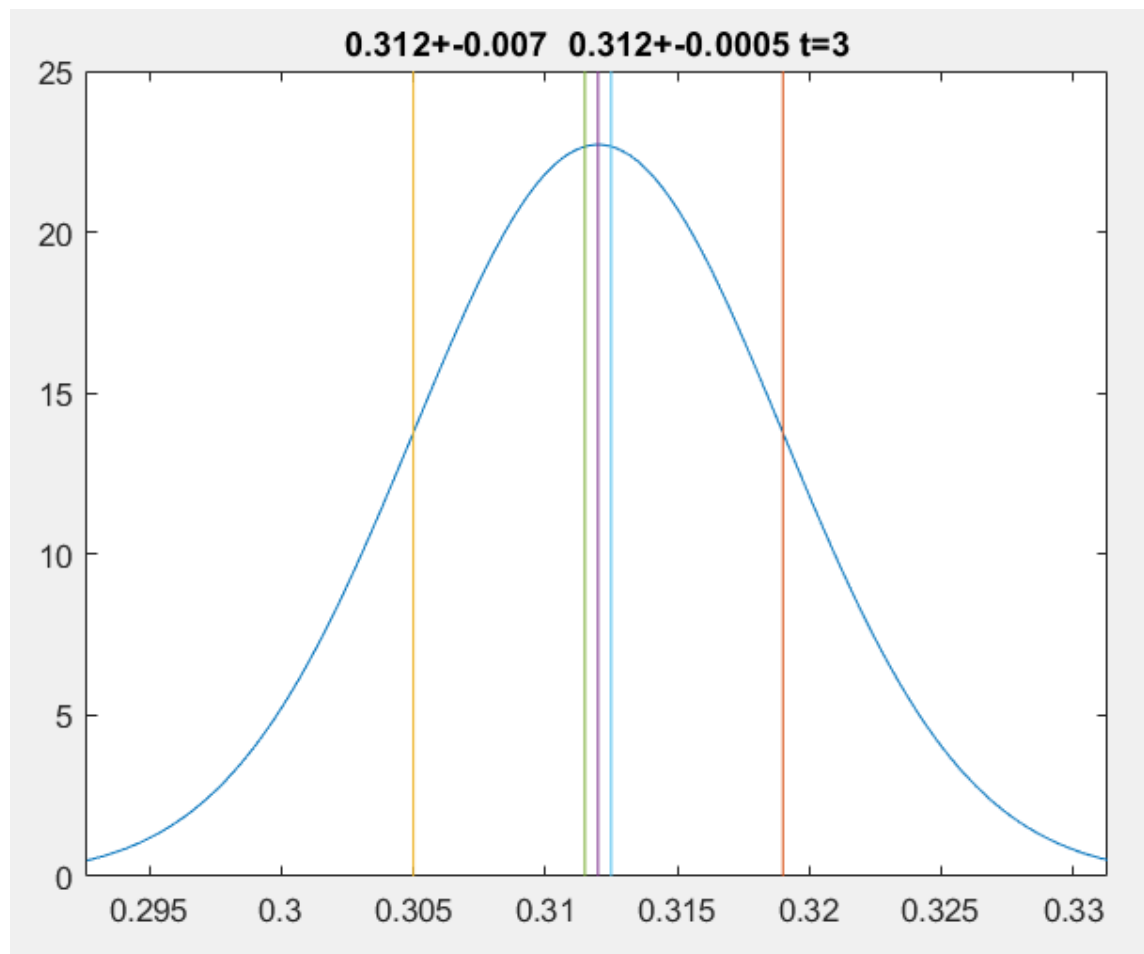
Nie ma sprzeczności 0.3¹2



Cyfra 2 jest poprawna

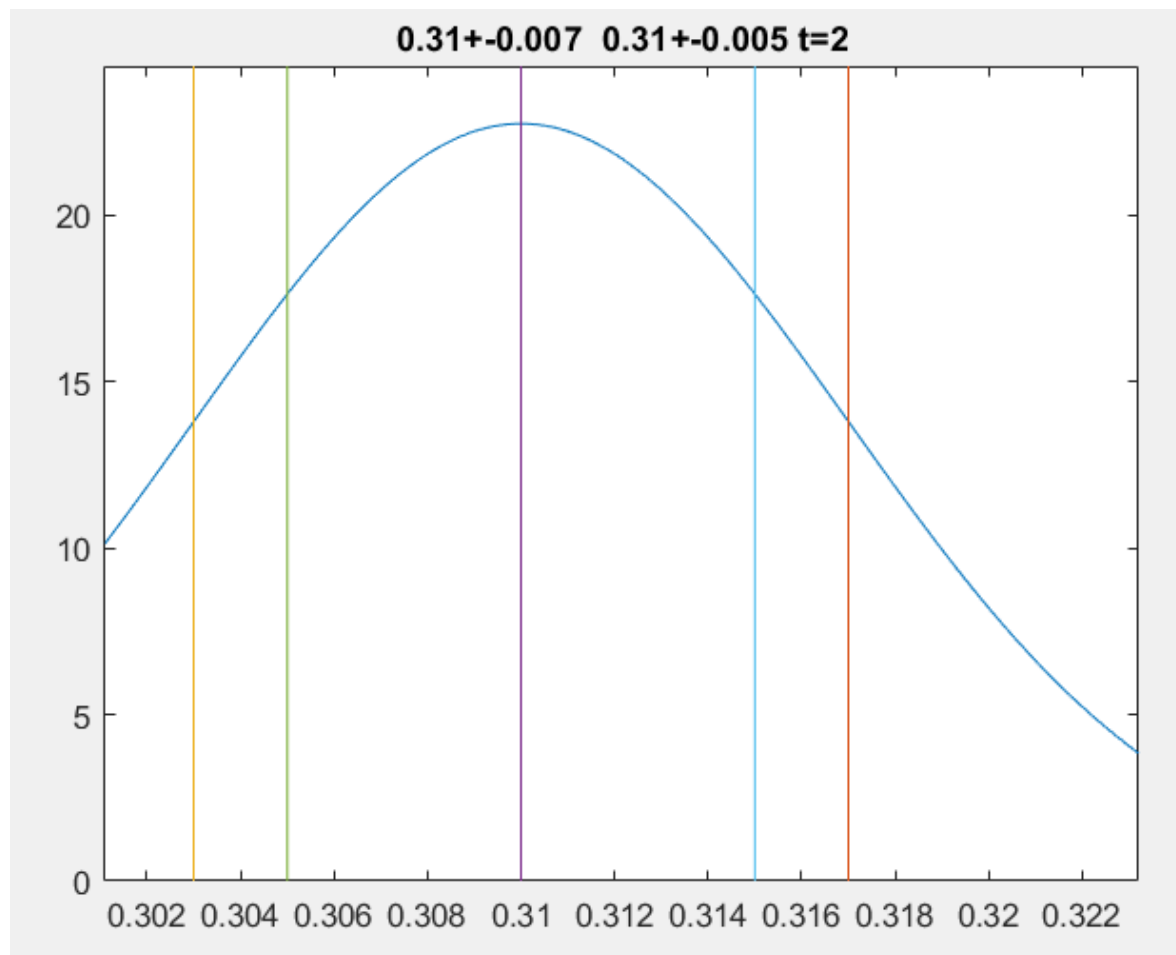
sprzeczność

0.312



Cyfra 1 jest poprawna

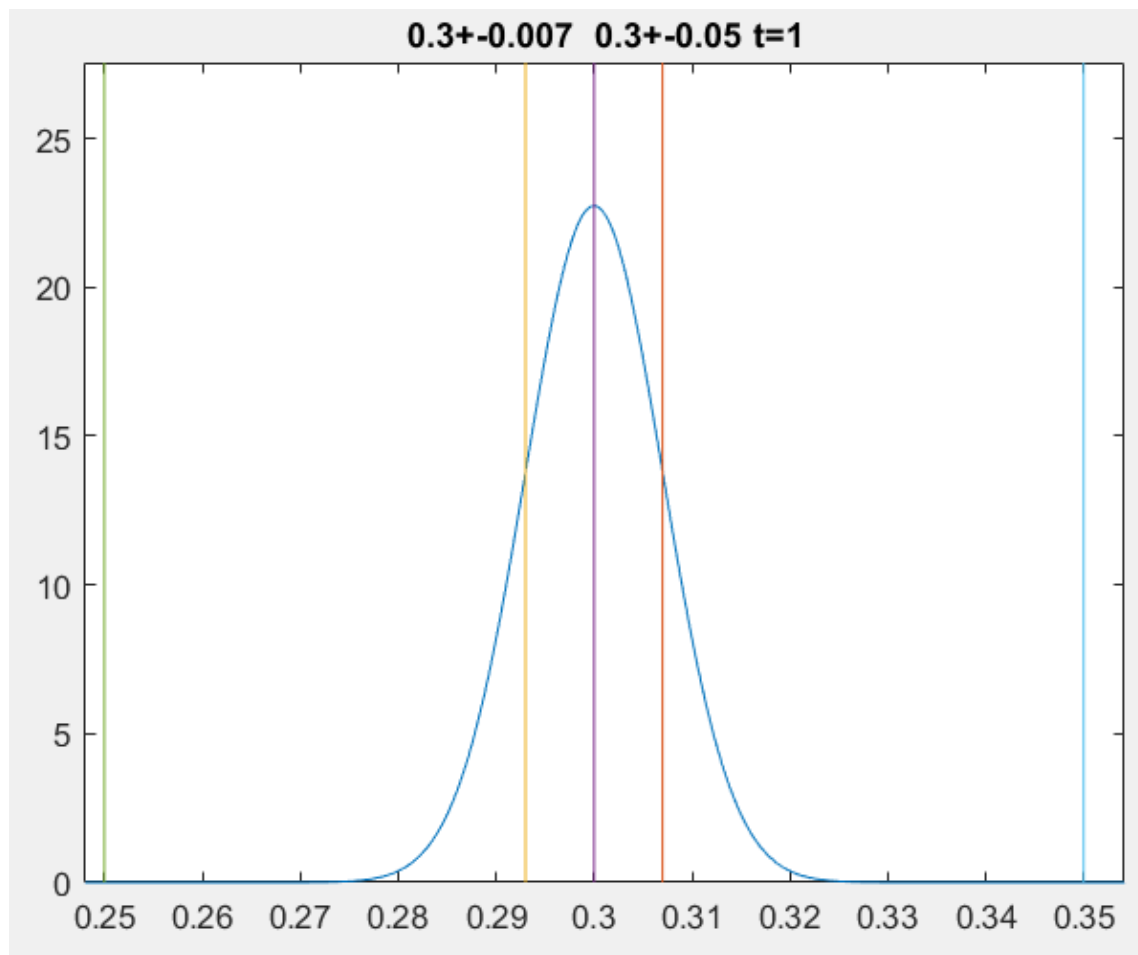
sprzeczność 0.3¹2



Cyfra 3 jest poprawna

Nie ma sprzeczności

0.**3**12



Błędy przypadkowe w pomiarach pośrednich

22-02

Założmy, że chcemy wyznaczyć pewną wielkość fizyczną **A**, ale **nie możemy jej zmierzyć bezpośrednio**. Wiemy natomiast, że jest ona związana z **K** innymi wielkościami fizycznymi X_1, X_2, \dots, X_K , które można już zmierzyć bezpośrednio, następującą zależnością:

$$A = f(X_1, X_2, \dots, X_K)$$

Po wykonaniu pomiarów mamy wyniki i błędy pomiarowe

$$\bar{X}_1 \pm \Delta \bar{X}_1$$

$$\bar{X}_2 \pm \Delta \bar{X}_2$$

.....

.....

$$\bar{X}_K \pm \Delta \bar{X}_K$$

Wynikową wartość wielkości A

$$\bar{A} = f(\bar{X}_1, \bar{X}_2, \dots, \bar{X}_K)$$

Przenoszenie się błędów

Podstawowymi operacjami arytmetycznymi wykonywanymi przez maszynę cyfrową są **dodawanie, odejmowanie, mnożenie, dzielenie**. Z tymi operacjami są związane błędy wynikające z niedokładności składników/czynników (operandów) i błędów zaokrągleń.

$x_1/\Delta x_1, x_2/\Delta x_2, \dots, x_n/\Delta x_n$ – operandy/błędy operandów

W ogólnym przypadku wyznaczona (obliczona) wielkość

$$y = f(x_1, x_2, \dots, x_n)$$

jest pewną funkcją wielu niezależnych zmiennych x_i , z których każda jest obciążona błędem Δx_i .

Ze wzoru Taylora wynika, że

$$\begin{aligned} y + \Delta y &= f(x_1 + \Delta x_1, x_2 + \Delta x_2, \dots, x_n + \Delta x_n) \\ &= f(x_1, x_2, \dots, x_n) + \sum_{i=1}^n \frac{\partial f}{\partial x_i} \Delta x_i + \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j} \Delta x_i \Delta x_j + \dots \end{aligned}$$

Przenoszenie się błędów (cd)

Jeśli założyć, że błędy Δx_i są małe, to

$$\Delta y \approx \sum_{i=1}^n \frac{\partial f}{\partial x_i} \Delta x_i$$

$$\Delta y \lesssim \sum_{i=1}^n \left| \frac{\partial f}{\partial x_i} \right| |\Delta x_i|$$

Jest to ogólny wzór dla szacowania błędów (przenoszenia się błędów).

Przenoszenie się błędów (cd)

Ze wzoru wynika, że dla sumy/różnicy dwóch wielkości x_1 i x_2 obarczonych błędami Δx_1 i Δx_2 otrzymujemy

$$y = x_1 \pm x_2$$

$$\Delta y = \Delta x_1 \pm \Delta x_2$$

Ponieważ błędy mogą być dodatnie bądź ujemne, więc przy szacowaniu błędu sumy musimy wziąć pod uwagę ich moduły, czyli błąd bezwzględny y wynosi

$$\Delta y \leq |\Delta x_1| + |\Delta x_2|$$

Przenoszenie się błędów (cd)

Błąd względny przy dodawaniu

$$\frac{\Delta y}{y} = \frac{|\Delta x_1| + |\Delta x_2|}{x_1 + x_2}$$

Błąd względny przy odejmowaniu

$$\frac{\Delta y}{y} = \frac{|\Delta x_1| + |\Delta x_2|}{x_1 - x_2}$$

Jeśli $x_1 \approx x_2$, to błąd względny może być dużo większy niż błędy względne obu składników!

Przykład: $\log \frac{a}{b} = \log a - \log b$, $a = 2.5000$ i $b = 2.4999$.

$$\log \frac{a}{b} = 0.39794 - 0.39792 = 0.00002,$$

Żadna cyfra wyniku nie jest poprawna, gdyż powinno być 0.000017372!

Problem: Odejmowanie bliskich sobie liczb, to główne źródło błędów.

Błąd maksymalny i standardowy

Dla $y = \sum_{i=1}^n x_i$ błąd wynosi $\Delta y = \sum_{i=1}^n \Delta x_i$.

Błąd sumy zależy liniowo od liczby składników. Jeśli $\Delta x_i = 0.5$, to suma 10^4 wyrazów będzie obciążona błędem 5×10^3 !

W praktyce nie wszystkie błędy mają taki sam znak, bo zaokrąglanie wprowadza błędy dodatnie i ujemne, które się kompensują.

W przypadku dużej liczby zmiennych błąd wielkości złożonej szacujemy przez podanie tzw. błędu standardowego (statystycznego).

Takie postępowanie jest uzasadnione, jeśli możemy traktować błędy poszczególnych zmiennych jako zmienne losowe niezależne o rozkładzie normalnym.

Błąd maksymalny i statystyczny (cd)

Tw. Załóżmy, że błędy Δx_i , $i = 1, \dots, n$ są niezależnymi zmiennymi losowymi o wartości oczekiwanej równej zero i odchyleniach standardowych ε_i . Wtedy błąd standardowy dla $y = f(x_1, x_2, \dots, x_n)$ jest dany wzorem

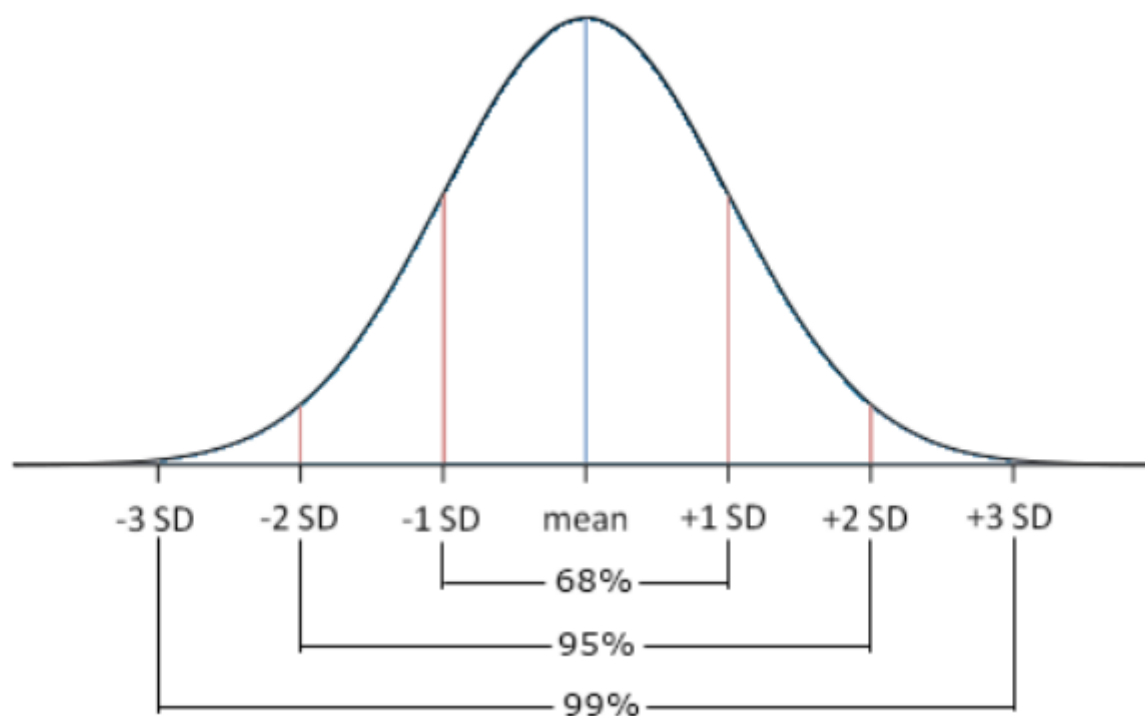
$$\left[\sum_{i=1}^n \left(\frac{\partial f}{\partial x_i} \right)^2 \varepsilon_i^2 \right]^{1/2}$$

Dla sumy n -składników o jednakowym błędzie ε błąd standardowy wynosi $\sqrt{\sum_{i=1}^n \varepsilon^2} = \varepsilon \sqrt{n}$.

Przy powyższych założeniach i $n \gg 1$ błąd sumy y ma w przybliżeniu rozkład normalny z wartością oczekiwaną 0 i odchyleniem standardowym $\sigma = \varepsilon \sqrt{n}$ i jest opisany wzorem

$$\rho(\Delta y) = \frac{1}{\sqrt{2\pi}\sigma} \exp(-\Delta y^2 / 2\sigma^2)$$

Błąd maksymalny i statystyczny (cd)



Obliczenie błędu kwadratowego $\Delta \bar{A}$

Jeżeli ΔX_i ($i = 1, 2, \dots, K$) są **średnimi błędami kwadratowymi** σ_{X_i} wartości średnich X_i , to otrzymujemy optymalnie znaleziony średni błąd kwadratowy z wyrażenia

$$\sigma_{\bar{A}} = \sqrt{\left[\left(\frac{\partial A}{\partial X_1} \right) \sigma_{\bar{X}_1} \right]^2 + \left[\left(\frac{\partial A}{\partial X_2} \right) \sigma_{\bar{X}_2} \right]^2 + \dots + \left[\left(\frac{\partial A}{\partial X_K} \right) \sigma_{\bar{X}_K} \right]^2}$$

Wtedy prawdopodobieństwo znalezienia wartości rzeczywistej A w przedziale

$$\langle \bar{A} - \sigma_{\bar{A}}, \bar{A} + \sigma_{\bar{A}} \rangle \text{ wynosi } 0,683.$$

Jeżeli błędy ΔX_i są **błędami granicznymi** (maksymalnymi) $\Delta X_{i \max}$, to w najmniej korzystnym przypadku otrzymujemy **błąd maksymalny pomiaru**

$$\Delta \bar{A}_{\text{gr}} = \left| \left(\frac{\partial A}{\partial X_1} \right) \Delta \bar{X}_{1 \max} \right| + \left| \left(\frac{\partial A}{\partial X_2} \right) \Delta \bar{X}_{2 \max} \right| + \dots + \left| \left(\frac{\partial A}{\partial X_K} \right) \Delta \bar{X}_{K \max} \right|$$

Wtedy prawdopodobieństwo znalezienia wartości rzeczywistej A w przedziale

$$\langle \bar{A} - \Delta \bar{A}_{\text{gr}}, \bar{A} + \Delta \bar{A}_{\text{gr}} \rangle \text{ wynosi } 0,999.$$

Przykład

Wahadło matematyczne o **długości $l = 100 \pm 1$ cm** posiada **okres wahań $T = 2,00 \pm 0,02$ s**. Wyznaczone błędy są średnimi błędami kwadratowymi. **Należy obliczyć przyspieszenie ziemskie.**

$$T = 2\pi \sqrt{\frac{l}{g}}$$



$$g = \frac{4 \pi^2 l}{T^2} = \frac{4 \pi^2 1m}{(2s)^2} = 9,87 \text{ m/s}^2$$

średni błąd kwadratowy przyspieszenia ziemskiego wynosi

$$\sigma_g = \sqrt{\left[\frac{4 \pi^2}{T^2} \sigma_l \right]^2 + \left[-\frac{4 \pi^2 l}{T^3} \sigma_T \right]^2}$$

$$\sigma_g = \sqrt{\left[\frac{4 \pi^2}{(2s)^2} 0,01m \right]^2 + \left[-\frac{4 \pi^2 1m}{(2s)^3} 0,02s \right]^2},$$

$$\sigma_g = \sqrt{\left[0,1 \text{ m/s}^2 \right]^2 + \left[0,2 \text{ m/s}^2 \right]^2} = \sqrt{0,01 + 0,04} \text{ m/s}^2,$$

$$\sigma_g = 0,224 \text{ m/s}^2$$

$$g = 9,87 \pm 0,23 \text{ m/s}^2$$

Metoda najmniejszych kwadratów

Założmy, że z doświadczenia uzyskaliśmy **n par wyników x_i, y_i** .

Szukamy zależność $y'_i = \bar{a}x_i + \bar{b}$

$$\sum_{i=1}^n (y'_i - y_i)^2 = \text{minimum}$$

$$\sum_{i=1}^n (\bar{a} x_i + \bar{b} - y_i)^2 = \text{minimum}$$

$$f = \sum_{i=1}^n (\bar{a} x_i + \bar{b} - y_i)^2$$

MNK

$$\sum_{i=1}^n 2 x_i (\bar{a} x_i + \bar{b} - y_i) = 0 ,$$

$$\sum_{i=1}^n 2 (\bar{a} x_i + \bar{b} - y_i) = 0 ,$$

$$2 \bar{a} \sum_{i=1}^n x_i^2 + 2 \bar{b} \sum_{i=1}^n x_i - 2 \sum_{i=1}^n x_i y_i = 0 ,$$

$$2 \bar{a} \sum_{i=1}^n x_i + 2 n \bar{b} - 2 \sum_{i=1}^n y_i = 0 .$$

MNK (cd)

$$\overline{a} = \frac{\sum_{i=1}^n x_i \sum_{i=1}^n y_i - n \sum_{i=1}^n (x_i y_i)}{\left(\sum_{i=1}^n x_i \right)^2 - n \sum_{i=1}^n x_i^2},$$

$$\overline{b} = \frac{\sum_{i=1}^n x_i \sum_{i=1}^n x_i y_i - \sum_{i=1}^n y_i \sum_{i=1}^n x_i^2}{\left(\sum_{i=1}^n x_i \right)^2 - n \sum_{i=1}^n x_i^2}.$$

MNK (cd)

Odchylenie standardowe wartości średnich

$$y' = \bar{a}x + \bar{b} = A(x, a, b)$$

$$\frac{\partial y}{\partial \bar{a}} = x,$$

$$\frac{\partial y}{\partial \bar{b}} = 1,$$

$$\frac{\partial y}{\partial x} = \bar{a},$$



$$\sigma_{y'} = \sqrt{[\bar{a}\sigma_x]^2 + [x\sigma_{\bar{a}}]^2 + [\sigma_{\bar{b}}]^2},$$

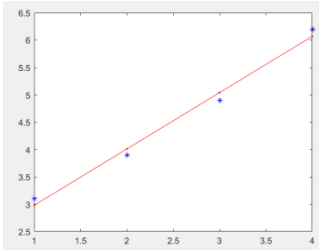
$$\sigma_x = 0$$

$$\sigma_{\bar{a}} = \sqrt{\frac{1}{n-2} \sum_{i=1}^n \varepsilon_i^2 \frac{n}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2}},$$

$$\sigma_{\bar{b}} = \sqrt{\frac{1}{n-2} \sum_{i=1}^n \varepsilon_i^2 \frac{\sum_{i=1}^n x_i^2}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2}},$$

$$\sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n y_i^2 - \bar{a} \sum_{i=1}^n x_i y_i - \bar{b} \sum_{i=1}^n y_i$$

Przykład



$$y' = \bar{a}x + \bar{b} = A(x, a, b)$$

$$a = 1.95$$

$$b = 1.03$$

$$S_a = 0.0793725393319376$$

$$S_b = 0.217370651192841$$

```
%mnkOS.m
x=[1 2 3 4];
y=[1.1 1.9 2.9 4.2]+2;
plot(x,y,'*b');
A =[1 x(1); 1 x(2); 1 x(3); 1 x(4)];
b=y;
l=A\b';
%The line then has the equation
a=l(2);b=l(1);
g=a.*x+b;
hold on;
plot(x,g,'.-r');
```

```
e=abs(y-g);
n=4;
Sa=sqrt(1/(n-2)*sum(e.*e)*n/(n*sum(x.*x)-sum(x)^2));
Sb=sqrt(1/(n-2)*sum(e.*e)*sum(x.*x)/(n*sum(x.*x)-sum(x)^2));
sume2=sum(y.*y)-a*sum(x.*y)-b*sum(y);
Sy=@(x)sqrt((x.*Sa).^2+(1*Sb).^2);
[x' y' g' Sy(x')]
```

| | x | y | g | +/- | Sy |
|--|---|--------|--------|-----|--------|
| | 1 | 3.1000 | 2.9800 | | 0.2314 |
| | 2 | 3.9000 | 4.0100 | | 0.2692 |
| | 3 | 4.9000 | 5.0400 | | 0.3224 |
| | 4 | 6.2000 | 6.0700 | | 0.3848 |

Błędy iteracji

$$e^x = \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n \approx 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!}$$

$$a = e^{0.5} = 1.64972\dots$$

$$\tilde{a} = 1$$

$$\delta = 39.3837\%$$

$$e^x = \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n \approx 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!}$$

$$a = e^{0.5} = 1.64972\dots$$

$$\tilde{a} = 1 + 0.5 = 1.5$$

$$\delta = 9.02\%$$

$$\varepsilon = \frac{\tilde{a}_{n+1} - \tilde{a}_n}{\tilde{a}_{n+1}}$$

$$\varepsilon = \frac{1.5 - 1}{1.5} 100\% = 33.3\%$$

| n | e^x | δ | ε |
|-----|-------------|----------|---------------|
| 1 | 1 | 39.3 | |
| 2 | 1.5 | 9.02 | 33.3 |
| 3 | 1.625 | 1.44 | 7.69 |
| 4 | 1.645833333 | 0.175 | 1.27 |
| 5 | 1.648437500 | 0.0172 | 0.158 |
| 6 | 1.648697917 | 0.00142 | 0.0158 |

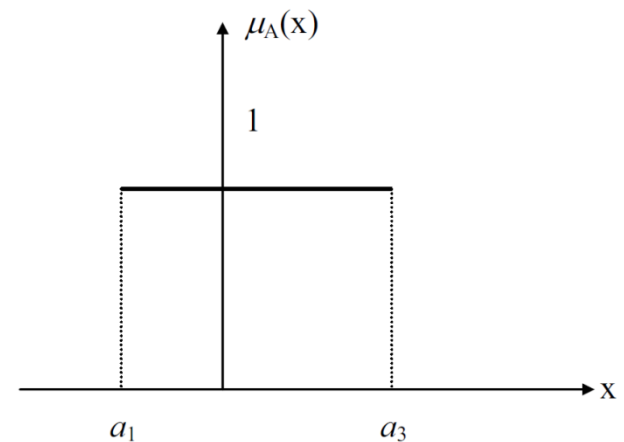
Działania na liczbach przybliżonych

- W przypadku dodawania lub odejmowania w wyniku pozostawiamy liczbę cyfr znaczących równą liczbie cyfr znaczących najmniej dokładnego składnika operacji (najmniejszą liczbę cyfr znaczących).
 - Ilość cyfr znaczących w ilorazie (w iloczynie) równa się ilości cyfr znaczących w tym jego czynniku, który zawiera najmniej cyfr znaczących.
 - W przypadku odwrotności liczby (o liczbie cyfr >2) ilość cyfr znaczących maleje o 1.
 - Przy dodawaniu, odejmowaniu i mnożeniu liczby przybliżonej z udziałem liczby dokładnej w wyniku podajemy taką ilość cyfr znaczących, jaką ma liczba przybliżona.
 - W przypadku pierwiastkowania liczby przybliżonej można przyjąć, że liczba cyfr pierwiastka jest równa liczbie cyfr liczby podpierwiastkowej (wskazówka praktyczna).
 - W przypadku logarytmowania liczby przybliżonej liczba cyfr znaczących pozostaje bez zmian
- a) $32,658 + 11,23 + 4,71 = 48,598 = 48,6$ $11 + 0,11 - 0,0011 = 11,1089$ zapis prawidłowy 11
 $9,01 + 21,5 + 12,456 = 42,961$ zapis prawidłowy 43,0
- b) $123,45 \cdot 12,3456 = 1524,06$; $9,32 \cdot 111 \cdot 0,038 = 39,31176 = 39$; $12,345^2 = 152,399$;
 $6,221$ (4 cyfry znaczące) $\cdot 5,2$ (2 cyfry znaczące) $= 32,3492 = 32$ (2 cyfry znaczące)
- Lub $15,5$ (3 cyfry znaczące) $\cdot 27,3$ (3 cyfry znaczące) $\cdot 5,4$ (2 cyfry znaczące) $= 2285,01 = 2,300 = 2,3 \cdot 10^3$ (2 cyfry znaczące).

Interwal. Definicja

$$A = [a_1, a_3] \quad a_1, a_3 \in \mathbb{R} \quad a_1 < a_3$$

$$\mu_A(x) = \begin{cases} 0, & x < a_1 \\ 1, & a_1 \leq x \leq a_3 \\ 0, & x > a_3 \end{cases}$$



Punkt jeżeli $a_1 = a_3$

Operacje na interwałach

$$[x_1, x_2] + [y_1, y_2] = [x_1 + y_1, x_2 + y_2]$$

$$[x_1, x_2] - [y_1, y_2] = [x_1 - y_2, x_2 - y_1]$$

$$[1,3] + [4,5] = [5,8], \quad 1 + [-3,2] = [-2,3]$$

$$[2,5] - [0,2] = [0,5], \quad [-2,-1] - [3,4] = [-6,-4]$$

Operacje na interwałach (cd)

$$[x_1, x_2] \cdot [y_1, y_2] = [\min\{x_1 y_1, x_1 y_2, x_2 y_1, x_2 y_2\}, \max\{x_1 y_1, x_1 y_2, x_2 y_1, x_2 y_2\}]$$

$$[1, 2] \cdot [3, 4] = [3, 8],$$

$$[-2, -1] \cdot [0, 2] = [-4, 0],$$

$$[-1, 2] \cdot [-1, 2] = [-2, 4]$$

Operacje na interwałach (cd)

$$\frac{[x_1, x_2]}{[y_1, y_2]} = [x_1, x_2] \cdot \frac{1}{[y_1, y_2]},$$

$$\left\{ \begin{array}{ll} \frac{1}{[y_1, y_2]} = \left[\frac{1}{y_2}, \frac{1}{y_1} \right] & \text{if } 0 \notin [y_1, y_2] \\ \frac{1}{[y_1, 0]} = \left[-\infty, \frac{1}{y_1} \right] \\ \frac{1}{[0, y_2]} = \left[\frac{1}{y_2}, \infty \right] \\ \frac{1}{[y_1, y_2]} = \left[-\infty, \frac{1}{y_1} \right] \cup \left[\frac{1}{y_2}, \infty \right] \subseteq [-\infty, \infty] & \text{if } 0 \in (y_1, y_2) \end{array} \right.$$

$[1,3]/[-1/1]$ nie jest zdefiniowane; $[-2,4]/[1,2] = [-2,4];$

$2/[1,2] = [1,2];$

$[2,4]/[-2,-1] = [-4,-1].$

Operacje na interwałach (cd)

potęgowanie x^y jest możliwe w jednym z przypadków

$$1^0 \quad \underline{x} > 0$$

$$2^0 \quad \underline{x} \geq 0 \wedge \underline{y} > 0$$

$$3^0 \quad 0 \notin x, y \leq 0 \text{ - całkowite}$$

$$4^0 \quad y > 0 \text{ - całkowite}$$

$$\text{np. } [2,3]^2 = [4,9]; \quad [-1,2]^2 = [2,4]$$

Własności działań na interwałach

1. $x + y = y + x$; $x \cdot y = y \cdot x$ (przemienność)
2. $(x + y) + z = x + (y + z)$; $(x \cdot y) \cdot z = x \cdot (y \cdot z)$ (łączność)
3. $x + 0 = 0 + x = x$; $1 \cdot x = x \cdot 1 = x$ (el. neutralny)
4. $x - y = x + (-y) = -y + x$
5. $x / y = x \cdot y^{-1} = y^{-1} \cdot x$
6. $-(x - y) = y - x$
7. $x(-y) = (-x) \cdot y = -x \cdot y$
8. $(-x) \cdot (-y) = x \cdot y$
9. $x - (y \pm z) = (x - y) \mp z$
10. $x \cdot (y \pm z) \subseteq x \cdot y \pm x \cdot z$
11. $(x \pm y) \cdot z \subseteq x \cdot y \pm y \cdot z$
12. $x - y \subseteq (x + y) - (y + z)$
13. $x / y \subseteq (x \cdot z) / (y \cdot z)$
14. $0 \in x - x$
15. $1 \in x / x$

Przykłady

$$[-1,1] \cdot ([0,1] + [-1,0]) = [-1,1] \subseteq [-2,2] = [-1,1] \cdot [0,1] + [-1,1] \cdot [-1,1] \cdot [-1,0]$$

$$1 \in [\frac{1}{2}, 2] = [1,2]/[1,2]$$

Przenoszenie się błędów

Niech $x_1 = 2.31 \pm 0.02$, $x_2 = 1.42 \pm 0.03$

Obliczmy różnicę:

Arytmetyka interwałów

$$x_1 - x_2 : 2.33 - 1.39 = 0.94, \quad 2.29 - 1.45 = 0.84,$$

czyli:

$$x_1 - x_2 = 0.89 \pm 0.05.$$

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

Zatem:

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

gdzie:

$\varepsilon_1, \varepsilon_2$ są maksymalnymi błędami bezwzględnymi.

Błąd bezwzględny sumy i różnicy jest więc równy: $\varepsilon_1 + \varepsilon_2$.

Przenoszenie się błędów (cd)

W przypadku obliczania iloczynu lub ilorazu przenoszenie błędów przedstawimy za pomocą **błędów względnych**.

Niech r będzie rzeczywistym błędem względnym, tzn.

$$\tilde{x} = x + rx = x(1 + r).$$

Weźmy:

$$\tilde{x}_1 = x_1(1 + r_1), \tilde{x}_2 = x_2(1 + r_2).$$

Wówczas:

$$\tilde{x}_1 \tilde{x}_2 = x_1 x_2 (1 + r_1)(1 + r_2), \quad \frac{\tilde{x}_1}{\tilde{x}_2} = \frac{x_1(1 + r_1)}{x_2(1 + r_2)}.$$

Błąd względny iloczynu jest zatem równy:

$$(1 + r_1)(1 + r_2) - 1 = 1 + r_1 + r_2 + r_1 r_2 - 1 \approx r_1 + r_2,$$

jeśli tylko $|r_1| \ll 1, |r_2| \ll 1$.

Iloraz:

$$\frac{(1 + r_1)}{(1 + r_2)} - 1 = \frac{1 + r_1 - 1 - r_2}{1 + r_2} = \frac{r_1 - r_2}{1 + r_2} \approx r_1 - r_2, \text{ jeśli } |r_2| \ll 1.$$

$$x = \tilde{x} \pm \varepsilon$$

$$r = \frac{\varepsilon}{x}$$

$$\varepsilon = rx$$

$$\tilde{x} = x + \varepsilon = x + rx = x(1 + r)$$

Uwarunkowanie zadania numerycznego

Małe zaburzenia danych powinny dawać małe zaburzenia wyniku, nie znajduje potwierdzenia nawet w prostych przypadkach. Umiejętność oceny jakościowego wpływu zaburzenia danych na wynik jest podstawą w obliczeniach numerycznych. **Wrażliwość rozwiązania** na dane początkowe określa tzw. **uwarunkowanie zadania numerycznego**.

Współczynniki uwarunkowania

- **uwarunkowanie względne:** jak względne zaburzenie danych wpływa na błąd względny wyniku:

$$\frac{\|f(x) - f(\tilde{x})\|}{\|f(x)\|} \leq \text{cond}_{rel}(f, x) \cdot \frac{\|x - \tilde{x}\|}{\|x\|}.$$

Najmniejszy mnożnik $\text{cond}_{rel}(f, x)$ spełniający powyższą nierówność nazywamy **współczynnikiem uwarunkowania (względnego)** zadania obliczenia $f(x)$ dla danego x .

- **uwarunkowanie bezwzględne:** jak bezwzględne zaburzenie danych wpływa na błąd bezwzględny wyniku:

$$\|f(x) - f(\tilde{x})\| \leq \text{cond}_{abs}(f, x) \cdot \|x - \tilde{x}\|.$$

Najmniejszy mnożnik $\text{cond}_{abs}(f, x)$ spełniający powyższą nierówność nazywamy **współczynnikiem uwarunkowania (bezwzględnego)** zadania obliczenia $f(x)$ dla danego x .

Wskaźnik uwarunkowania

Mówimy, że zadanie $f(x)$ jest

- *dobrze uwarunkowane* w punkcie x , gdy $\text{cond}(f, x) \approx 1$,
- *źle uwarunkowane* w punkcie x , gdy $\text{cond}(f, x) \gg 1$,
- *źle postawione* w punkcie x , gdy $\text{cond}(f, x) = +\infty$.

$$\begin{array}{l} x, f(x) \\ \tilde{x}, f(\tilde{x}) \end{array}$$

$$w = \frac{|\delta_f|}{|\delta_x|} = \frac{|(f(x) - f(\tilde{x})) / f(x)|}{|(x - \tilde{x}) / x|}$$

Przykład

$$x = 5cm$$

$$\tilde{x} = 4cm$$

$$\Delta = x - \tilde{x} = 1cm$$

$$\delta = 1 / 5 * 100\% = 20\%$$

$$f(x) = \pi x$$

$$g(x) = \pi \frac{x^2}{4}$$



$$f(x) = \pi * 5 \approx 15.708$$

$$f(\tilde{x}) = \pi * 4 \approx 12.5664$$

$$w = 1$$

$$g(x) = \pi \frac{5^2}{4} \approx 19.6350$$

$$g(\tilde{x}) = \pi \frac{4^2}{4} \approx 12.5664$$

$$w = 1.8$$



Problem jest niestabilny gdy $w \gg 1$

$$w = \frac{|\delta_f|}{|\delta_x|} = \frac{|(f(x) - f(\tilde{x})) / f(x)|}{|(x - \tilde{x}) / x|}$$

Wskaźnik uwarunkowania(dalszy ciąg)

$$x, f(x)$$

$$\tilde{x} = x + h, f(\tilde{x})$$

$$f(\tilde{x}) - f(x) = f(x + h) - f(x) \approx h \cdot f'(x)$$

$$w = \frac{|\delta_f|}{|\delta_x|} = \frac{|(f(x) - f(\tilde{x})) / f(x)|}{|(x - \tilde{x}) / x|} \approx$$

$$\left| \frac{h \cdot f'(x) / f(x)}{h / x} \right| = \left| \frac{x \cdot f'(x)}{f(x)} \right|$$

Przykład

$$f(x) = e^x$$

$$w \approx \left| \frac{x \cdot f'(x)}{f(x)} \right| = |x|$$

```
>> exp(1)
```

```
ans =
```

```
2.7183
```

```
>> exp(1.1)
```

```
ans =
```

```
3.0042
```

```
>> exp(4)
```

```
ans =
```

```
54.5982
```

```
>> exp(4.1)
```

```
ans =
```

```
60.3403
```

Przykład

$$f(x) = \cos(x)$$

$$w \approx \left| \frac{x \cdot f'(x)}{f(x)} \right| = |x \cdot \tan(x)|$$

$$\tan(\pi / 2) = \infty$$

```
>> cos(1.57079)
```

```
ans =
```

```
6.3268e-06
```

```
>> cos(1.57078)
```

```
ans =
```

```
1.6327e-05
```

```
>> (1.57079-1.57078)/1.57079
```

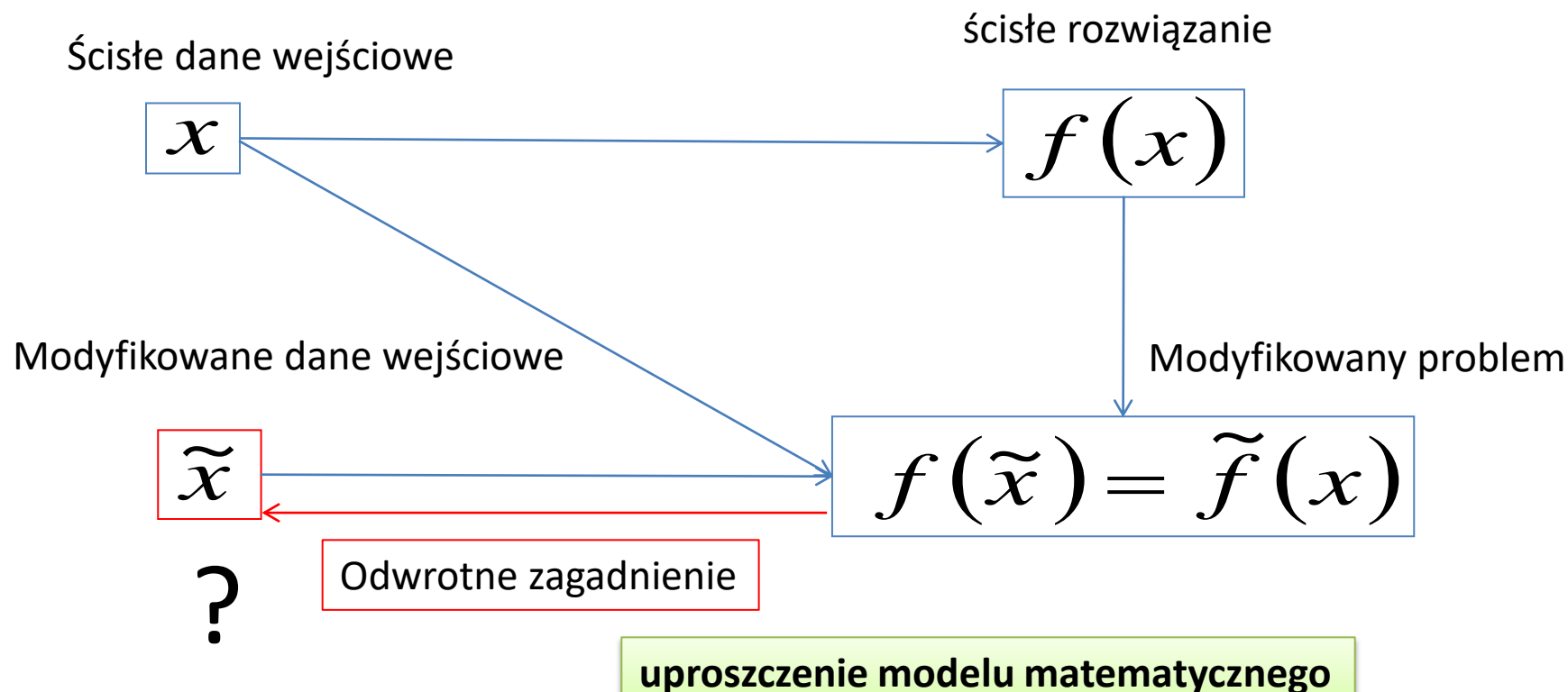
```
ans = 6.3662e-06
```

```
>> (cos(1.57079)-cos(1.57078))/cos(1.57079)
```

```
ans = -1.5806
```

```
w=2.4828e+05
```

Wsteczna analiza błędów



Analiza błędów obliczeniowych jest zwykle bardzo złożona. Dlatego przeprowadza się wsteczną analizę błędów.

Założmy, że rozwiązanie które mamy jest ścisłym rozwiązaniem dla zmodyfikowanego problemu. O ile trzeba zmodyfikować dane wejściowe, żeby otrzymać taki sam wynik dla problemu oryginalnego?

Przykład

$$f(x) = e^x = \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n} \right)^n = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

Modyfikowany problem

$$\tilde{f}(x) = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!}$$

$$x = 1$$

ściśle rozwiązanie

$$f(x) = 2.718282$$

$$\tilde{f}(x) = 2.666667$$

$$\tilde{f}(x) - f(x) = -0.051615$$

Modyfikowane dane
wejściowe

$$\tilde{x} = \log(\tilde{f}(x)) = \log(2.666667) = 0.980829$$

Wsteczny błąd

$$\tilde{x} - x = -0.019171$$

```
>> exp(0.980829)=2.6667
```

Wskaźnik uwarunkowania dla zadań i algorytmów

Przykład: Należy wyznaczyć x , y oraz $x + y$ z układu równań

$$x + \alpha y = 1$$

$$\alpha x + y = 0$$

$x = 1/(1 - \alpha^2)$ i wskaźnik uwarunkowania dla x wynosi

$$\frac{\Delta x/x}{\Delta \alpha/\alpha} = \frac{2\alpha^2}{1 - \alpha^2}$$

Algorytm jest dobrze uwarunkowany dla $\alpha \ll 1$ i $\alpha \gg 1$, źle dla $\alpha \approx 1$.

Obliczanie $y = -\alpha/(1 - \alpha^2)$ jest źle uwarunkowane dla $\alpha^2 \approx 1$.

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

jest dobrze uwarunkowane dla $\alpha \approx 1$.

Złe uwarunkowanie algorytmu można często zmniejszyć przez przeprowadzenie obliczeń w podwyższonej precyzji. Wskaźniki uwarunkowania algorytmu i zadania są od siebie wzajemnie niezależne.