Numerical computation

Computational mathematics

Theory of computation

Calculating mathematics

Numerical analysis

Curriculum/ test

Theory + Problems	Laboratory
32	32

http://en.wikipedia.org/wiki/Mathematics#Computational_mathematics

Kochegurova Elena (Кочегурова Елена Алексеевна)

CALCULATING AND ALGORITHMS ERROR (TOLERANCE)

1. The errors in the calculations.
2. Stability and complexity of the algorithm.
3. The classification errors.
4. Absolute and relative errors.
4. Direct and inverse problems of the theory of errors.
5. Unstable algorithms.
6. Features of computer arithmetic.

PROBLEM OF CLASSICAL MATHEMATICS

- ✓ To determine the existence and uniqueness of solutions.

Minuses (lows) :

- ❑ Impossibility solve this problem;
- ❑ Impossibility practical using the solution
- ❑(obtained solution - cumbersome (lenght).

Problems of numerical Mathematics

- ✓ Find a solution to the required accuracy. ($\varepsilon=10^{-3} - 10^{-6}$)

Highs :

- ❑ Ability to obtain solutions with different accuracies to get the result

APPROXIMATE CALCULATION

7/3 2.33333…

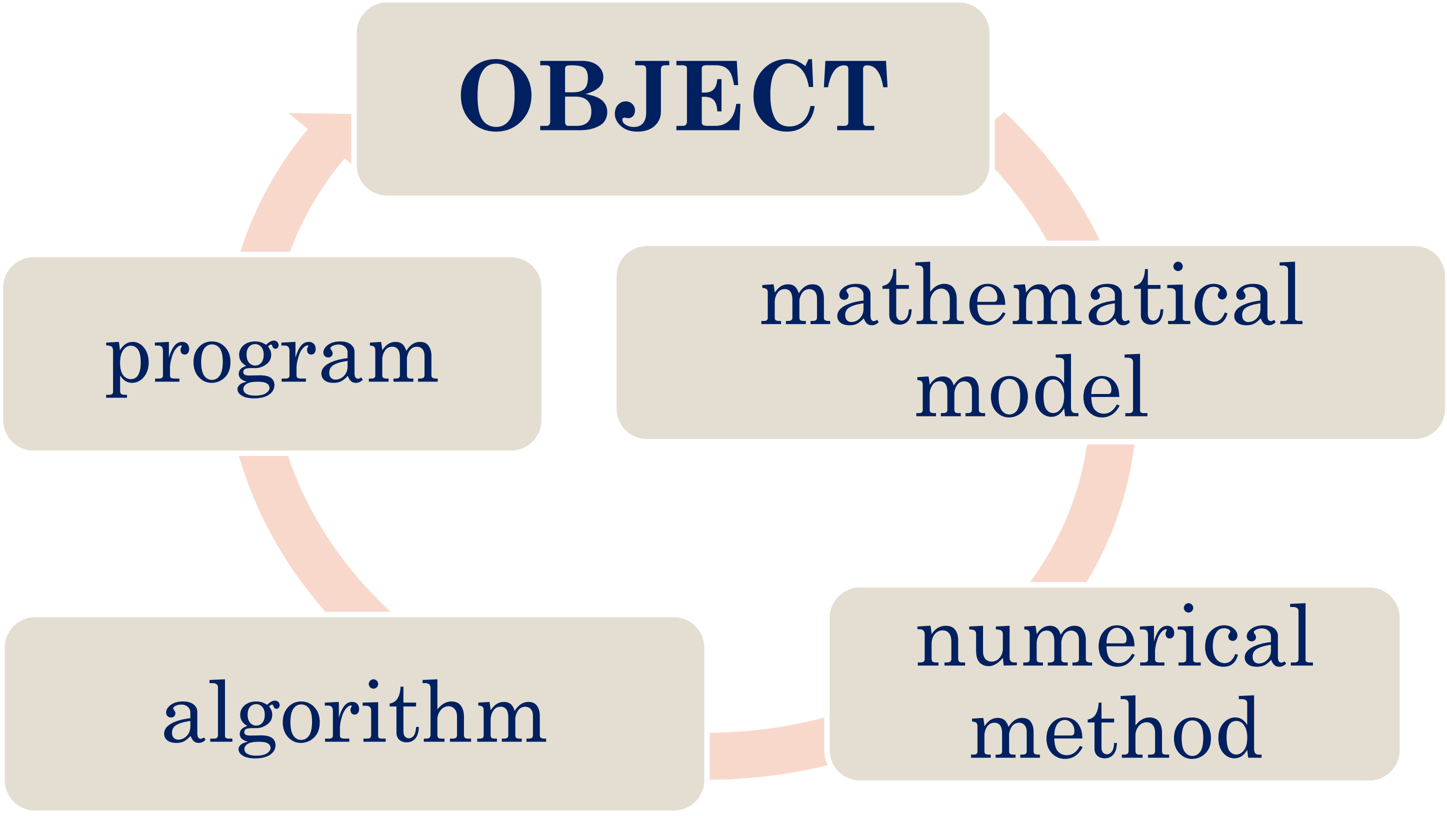
Mathematica (company Wolfram Research)

Maple (company Waterloo Maple Inc)

MatLab (company The MathWorks)

MathCAD (company MathSoft Inc).

COMPUTATIONAL EXPERIMENT



REQUIREMENTS FOR COMPUTING (NUMERICAL) METHODS

- The adequacy of the discrete model of the original mathematical problem :
  - stability,
  - convergence,
  - correctness.

- The possibility of realization the discrete model on a computer.

The Concept of
NUMERICAL METHODS STABILITY

Definition 1:

The words "stable algorithm" means that more accurate input data can improve the result.

Definition 2:

The *stability of the algorithm* means that small deviations in the input data correspond to small deviations in the solution.

The Concept of Numerical Methods Convergence

Definition 1:

Convergence means the **<u>closeness</u>** (aboutness, proximity) of the resulting numerical solution to the true solution

Definition 2:

Convergence of the numerical method (algorithm) means the ability of the method to obtain the exact solution after a finite number of steps, with any desired accuracy for any initial approximation

THE CORRECTNESS OF NUMERICAL METHOD

Definition 1:

The task is set correctly, if for any values of the initial data its solution:

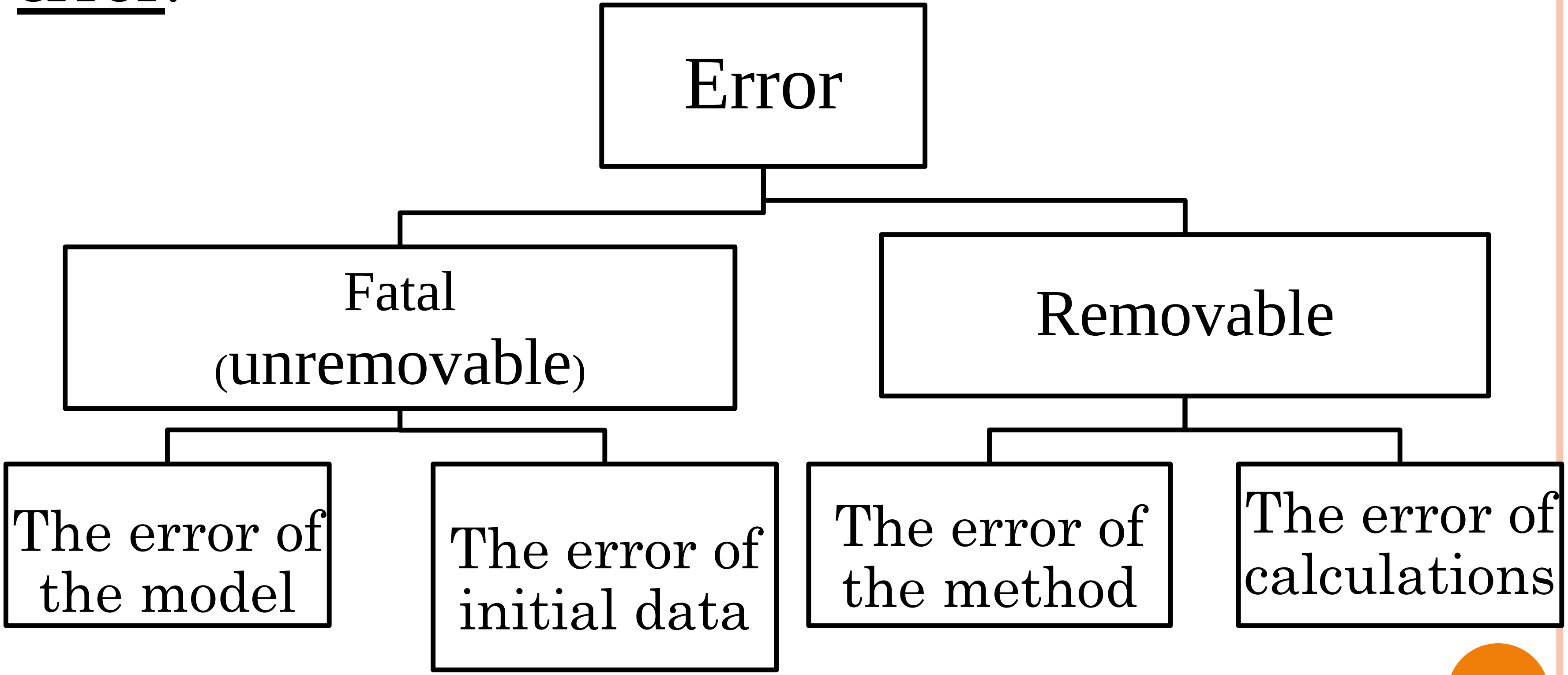
- exists;
- unique;
- stable.

Sometimes for solving correct problem can be use unstable method for its solution.

CONCEPT AND CLASSIFICATION OF ERROR

Definition :

The deviation from the true solution is an approximate **error**.



GLOBAL (FULL) ERROR

Full (total) error of the numerical solution includes :

- ***Fatal error*** connect with the error of the task and the inaccuracy of the initial data;

- ***Removable*** error includes error method of solving the task and calculation errors.

GLOBAL (FULL) ERROR

Removable error can be reduced by

- *choosing a more perfect (exact) method*
- *increasing in the bit numbers (digit capacity) of computer.*

Characteristics of accuracy of the solution of the problem is *absolute and relative* errors.

Absolute error

Approximate number ***X**** is a number just a little different from the exact X and replacing it in the calculations.

Definition :

Let ***X*** *- the exact solution,*

X* *- approximate solution.*

Then the ***absolute error*** of the approximate number X* - called value **Δ**, which is a limit of the difference

$$|\mathbf{X} - \mathbf{X^*}| \leq \mathbf{\Delta}$$

ABSOLUTE ERROR RECORDS

$$|\mathbf{X} - \mathbf{X^*}| \leq \Delta$$

$$\mathbf{X^*} - \Delta \leq \mathbf{X} \leq \mathbf{X^*} + \Delta$$

mathematical estimates of the error

$$\mathbf{X} = \mathbf{X^*} \pm \Delta$$

error of physical systems and instrumentation

ABSOLUTE ERROR

Example 1.1.

Let the length of the interval L = 100 cm was measured with an accuracy up to 0.5 cm.

Then write

$$L = 100 \text{ cm} \pm 0.5 \text{ cm}$$

Here, the absolute error **Δ** = 0.5 cm, and the exact value of the length L of the segment is contained within

$$99.5 \text{ cm} < L < 100.5 \text{ cm}$$

In the measurements resulting record, it usually indicates the absolute error.

ABSOLUTE ERROR

Example 1.2.

To determine the absolute error of the approximate number ***X* =3.14***, which would replace **π**.

Solution:

As $3.14 < \pi < 3.15$, then modulus $|\mathbf{X^* - X}| < 0.01$;

Therefore, we can take $\Delta = 0.01$.

However, if we consider another representation of number **π**

$$3.14 < \pi < 3.142,$$

then we have a better estimate : $\Delta = 0.002$.

ABSOLUTE ERROR

Summary :

- There may be several values of the absolute error, each of which is determined by the boundaries of the approximate value of the number;

- Absolute accuracy is not sufficient (enough, adequative) characteristic of the accuracy of measurement or calculation.

RELATIVE ERROR

Definition :

The relative error of approximate number $\mathbf{X^*}$ – call the value of δ, defined by expression

$$\delta = \frac{\Delta}{\mathbf{X^*}}$$

Other form of record :

$$\mathbf{X^*}(1 - \delta) \leq \mathbf{X} \leq \mathbf{X^*}(1 + \delta)$$

RELATIVE ERROR

More popularly in engineering and technical applications to express the relative error as a percentage

- ✓ It is considered allowable error 3–5 % (in individual tasks to 10%)

RELATIVE ERROR

Example 1.3.

Let the number ***e*** is set by expression **e**=2.718±0.001

It is required to find the relative error of computation.

Solution.

According to a formula of absolute error

(X=X* ± Δ)

it will be obtained:

X=e; **X***=2.718; **Δ**=0.001.

We calculate the relative error :

$$\delta = \left(\frac{\Delta}{X^*}\right)*100\% = \left(\frac{0,001}{2,718}\right)*100\% \approx 0,036\%$$

RELATIVE ERROR

Example1.4.

Find the relative error of the measured lengths of the segments

а) L1 = 50.8 см ± 0.5 см

б) L2 = 3.6 см ± 0.5 см

Solution:

	а)	б)
X^*	50.8	3.6
Δ	0.5	0.5
δ	$\delta = \frac{0,5}{50,8} \cdot 100\% \approx 0,984\%$	$\delta = \frac{0,5}{3,6} \cdot 100\% \approx 13,888\%$

RELATIVE ERROR

Summary:

- At the same absolute error **Δ**=0.5,

the relative errors of measurement segments L1 and L2 *differ greatly*.

METHODS (way) OF REPRESENTATION OF REAL NUMBERS

Representation in shape (form) from the ***fixed -point***:

$$X^* = \alpha_1 \cdot \beta^{n} + \alpha_2 \cdot \beta^{n-1} + \ldots + \alpha_m \cdot \beta^{n-m+1}$$

α_1 – first significant figure;

β – base of numerical system (2,8,10,16) or base of positional system;

$0 \le \alpha_i \le \beta$ – numbers from a basis set.

METHODS (way) OF REPRESENTATION OF REAL NUMBERS

In the computer representation of numbers in the form of a *floating-point* most commonly used **:**

$$\mathbf{X^* = M \cdot \beta^p}$$

β – base of numerical system ;

p – order number (integral number positive, the negative or zero);

M – mantissa of number, $\beta^{-1} < M < 1$

METHODS (way) OF REPRESENTATION OF REAL NUMBERS

***Example* 1.5.**

Get decomposition of **2.718** in the form with the fixed- and floating- point.

With the fixed-point :

$$\mathbf{2.718 = 2 \cdot 10^{0} + 7 \cdot 10^{-1} + 1 \cdot 10^{-2} + 8 \cdot 10^{-3}}$$

$\beta=10; \alpha_1=2; \alpha_2=7; \alpha_3=1; \alpha_4=8;$ **n=0**.

With the floating-point :

$$\mathbf{2{,}718 = 0{,}2718 \cdot 10^{1}}$$

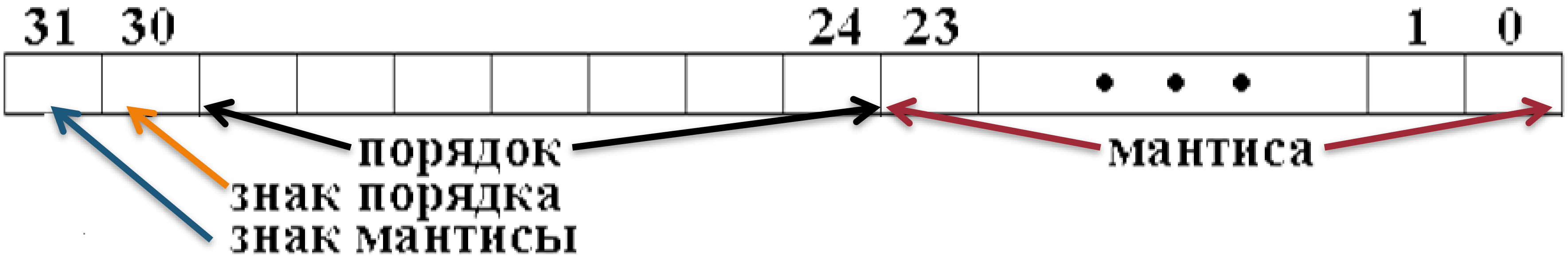
M=0.2718; p=1.

DIGIT GRID COMPUTER

Definition:

The digit grid of a computer is a number of the bits allocated to record the number.

For example, 32-bit grid:



The more bits in the computer, the greater the range of valid numbers.

Therefore less computation error.

CALCULATION ACCURACY

Accuracy of computation is defined by number of digits of the result credible

Trust characteristics to digits of result are :

- Significant,
- Valid,
- Dubious (uncertain) digits.

SIGNIFICANT, VALID, DUBIOUS DIGITS

Definition_:

Significant digit of number X* call all digits in its record, since the first nonzero at the left.

Definition_:

The significant figure $\boldsymbol{\alpha}_{\kappa}$ is considered ***valid*** (correct), if the inequality is executed:

$$\Delta \leq \omega \cdot \beta^{n-\kappa+1}$$

0.5≤ ω≤ 1 (most often in the calculations ω = 0.75)

otherwise– $\boldsymbol{\alpha}_{\kappa}$ – ***dubious*** figure.

SIGNIFICANT, VALID, DUBIOUS DIGITS

Example 1.6.

Determine the number of valid (correct) digits in the record e =2.718±0.001.

Solution. *Get decomposition with the fixed-point :*

$$\mathbf{X^* = 2.718 = 2 \cdot 10^{0} + 7 \cdot 10^{-1} + 1 \cdot 10^{-2} + 8 \cdot 10^{-3}}$$

$\beta=10$; $\alpha_1=2$; $\alpha_2=7$; $\alpha_3=1$; $\alpha_4=8$; **n=0**.

$$\Delta \le \omega \cdot \beta^{n-\kappa+1}$$

Absolute error $\Delta = 0.001$

with the fixed-point $\Delta=0.1 \cdot 10^{-2}$.

Select $\omega=0.75$.

SIGNIFICANT, VALID, DUBIOUS DIGITS

Get

$$\mathbf{0.1 \cdot 10^{-2} \leq 0.75 \cdot 10^{\,0-k+1}}$$

where ***k*** – unknown variable.

The same bases (10) and the number of mantissa (0,1<0,75) allow us to go to the inequality on the indicators:

$$\mathbf{-2 < 1-k}$$

Then **k ≤ 3 (**X* =**2.718)**

With the result that :

- *Correct digits of number are the three first significant figures, i.e.* **2.71**;
- Digit ***8*** – dubious.

COMMON FORMULA ERRORS

The general error of result is defined:

- values of separate (individual) errors,
- type of mathematical expression.

Rules of transformation errors :

- ✓ *The absolute error of the sum* of a finite number of approximate numbers doesn't exceed the sum of absolute errors of these numbers.
- ✓ *The relative error* of multiplication of a finite number of approximate numbers doesn't exceed the sum of the relative errors of these numbers.

COMMON FORMULA ERRORS

Let function is set $Z=f(x_1, x_2, \ldots, x_n)$

There are absolute (or relative) error argument
$\Delta x_1, \Delta x_2, \ldots, \Delta x_n$ ($\delta x_1, \delta x_2, \ldots, \delta x_n$).

Absolute error function :
(partial derivative)

$$\Delta Z = \sum_{i=1}^{n} \left| \frac{\partial Z}{\partial x_i} \right| \cdot \Delta x_i$$

Relative error :

$$\delta Z = \sum_{i=1}^{n} \left| x_i \cdot \frac{\partial}{\partial x_i} \ln Z \right| \cdot \delta x_i$$

COMMON FORMULA ERRORS

If a complex function Z depends on two arguments, i.e. **Z=f(x, y)**, then:

Absolute error	**Relative error**
$\Delta(x \pm y) = \Delta x + \Delta y$	$\delta(x \pm y) = \dfrac{x \cdot \delta x + y \cdot \delta y}{x \pm y}$
$\Delta(x \cdot y) = x \cdot \Delta y + y \cdot \Delta x$	$\delta(x \cdot y) = \delta x + \delta y$
$\Delta\left(\dfrac{x}{y}\right) = \dfrac{y \cdot \Delta x + x \cdot \Delta y}{y^2}$	$\delta\left(\dfrac{x}{y}\right) = \delta x + \delta y$
$\Delta(x^m) = m \cdot x^{m-1} \cdot \Delta x$	$\delta(x^m) = m \cdot \delta x$

COMMON FORMULA ERRORS

Example 1.7.

To calculate value of analytical expression and to evaluate absolute and relative errors of a composite function

$$Z = \frac{a \cdot b^3}{\sqrt{c}}$$

for given values

a = 0.643 ± 0.0005

b = 2.17 ± 0.002

c = 5.843 ± 0.001

COMMON FORMULA ERRORS

Solution.

Let's calculate value of approximate number of Z^*, substituting the values of its entering arguments a, b, c.

$$Z^* = \frac{0{,}643 \cdot 2{,}17^3}{5{,}843} = 27{,}1814$$

Absolute errors from a statement of the problem are equal

$\Delta a = 0{,}0005$; $\Delta b = 0{,}02$; $\Delta c = 0{,}001$.

e relative errors we will find from a formula: $\delta = \frac{\Delta}{X^*}$

$\delta a = 7{,}796 \cdot 10^{-3}$; $\delta b = 0{,}929$; $\delta c = 0{,}171$.

COMMON FORMULA ERRORS

For computation of errors of a composite function of **Z** we will use the formula obtained earlier:

$$\delta Z = \delta a + \delta(b^3) + \delta(\sqrt{c})$$

According to the formulas from the table of the *relative errors* we lead a formula to a look :

$$\delta Z = \delta a + 3 \cdot \delta(b) + \frac{1}{2} \cdot \delta(c) = 2{,}8385$$

For computation of *absolute error* of function **Z** we will use a formula:

$$\Delta Z = \frac{\delta \cdot X^*}{100\%} = \frac{2{,}8385 \cdot 27{,}1814}{100} = 0{,}7769$$

COMMON FORMULA ERRORS

<u>Пример 1.8</u>.
Вычислить значение аналитического выражения и оценить абсолютную и относительную погрешности сложной функции

$$Z = \frac{a^2 - b^2}{(a+b)^2} + h^2$$

при заданных значениях
$a = 0.643 \pm 0.0005$
$b = 2.17 \pm 0.002$
$c = 5.843 \pm 0.001$

COMMON FORMULA ERRORS

Решение.

Вычислим значение приближенного числа Z^*, подставив значения входящих в него аргументов a, b, h.

$$\mathbf{Z^* = 0{,}8307}$$

Абсолютные погрешности из условия задачи равны $\Delta a = 0{,}04$; $\Delta b = 0{,}02$; $\Delta h = 0{,}01$.

Относительные погрешности найдем как: $\delta = \dfrac{\Delta}{X^*}$

$\delta a = 3{,}5057\ \%$; $\delta b = 0{,}6337\ \%$; $\delta h = 0{,}8772\ \%$.

COMMON FORMULA ERRORS

В данной сложной функции Z^* основная математическая операция сложение.

Воспользуемся формулой для вычисления абсолютной погрешности, предварительно упростив первое слагаемое функции:

$$Z = \frac{a - b}{(a + b)} + h^2$$

$$\Delta Z = \frac{2 \cdot a \cdot (\Delta a + \Delta b)}{(a + b)^2} + 2 \cdot h \cdot \Delta h$$

Подставив численные значения абсолютных погрешностей, получим: $\Delta Z = 0{,}043$

Относительную погрешность функции пересчитаем по формуле $\delta = \frac{\Delta}{X^*} \cdot 100\%$ $\delta Z = 5.214\%$