

Введение в численные методы.

Представление чисел в ЭВМ.

Вычислительная погрешность. Обзор
инструментальных программных
средств.

Тема 1

Введение в численные методы

1.1 Математическое моделирование и использование вычислительной техники в решении прикладных задач.

Вычислительный эксперимент и его этапы.

Вычислительные задачи. **Корректность и обусловленность вычислительных задач.**

Вычислительные алгоритмы.

Численные методы. **Корректность и реализуемость численного метода**

Введение в численные методы

1.2 Источники и классификация погрешностей.

Неустраняемая погрешность. Погрешность метода.

Приближенные числа. Абсолютная и относительная погрешности. Значащие и верные цифры. Погрешности (относительные) арифметических операций. Погрешность функции одной и многих переменных. Обусловленность вычислительной задачи.

Представление чисел в ЭВМ. Понятия машинного эпсилон, машинной бесконечности, машинного нуля. Катастрофическая потеря точности.

1.3 Обзор инструментальных программных средств пакетов прикладных программ Mathematica, Mathcad и др.

Введение в численные методы

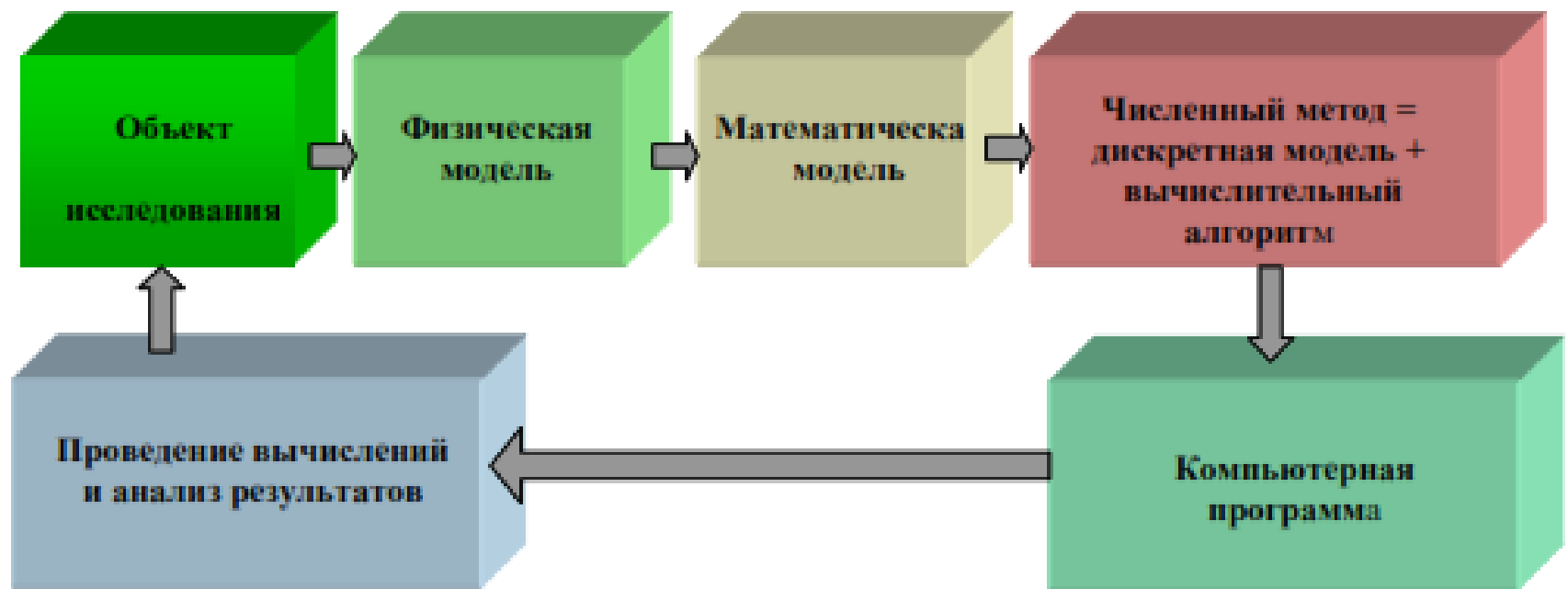


Схема вычислительного эксперимента

Введение в численные методы. Погрешности

На каждом этапе математического моделирования вносятся те или иные **погрешности**. Они обусловлены следующими причинами:

- a) математическое **описание задачи является упрощенным**;
- b) **недостаточно точно заданы исходные данные**, являющиеся, как правило, результатом проведенных экспериментов;
- c) всякий **численный метод является приближенным** в том смысле, что он не дает точного решения задачи, (для точного решения, например, может потребоваться неограниченно большое число арифметических операций);
- d) в самом процессе решения задачи **в ЭВМ при** вводе исходных данных, **выполнении арифметических операций** и при выводе результатов **производятся округления**.

Погрешности в решении задачи, обусловленные пунктами «a» и «b», называются **неустраняемыми**.

При переходе от математической модели к численному методу (пункт «c») возникает *погрешность метода*.

Построение численного метода для математической модели состоит из двух этапов – формулировки дискретной модели и разработки вычислительного алгоритма, позволяющего отыскать решение дискретной задачи.

Соответственно погрешность метода подразделяется на *погрешность дискретизации* или *погрешность аппроксимации* и на возникающую в процессе решения (пункт «d») *вычислительную погрешность* или *погрешность округления*.

Основы теории погрешностей

Выбранный приближенный метод реализуется неточно из-за **ошибок округления**, возникающих при вычислениях на реальном компьютере.

Изменение результатов вычислений, вызванное ошибками округления, называется **погрешностью округления** и **сильно зависит от числа значащих цифр, используемых для записи чисел в ЭВМ.**

Учет этого вида погрешности является наиболее сложным ввиду отсутствия ассоциативности и дистрибутивности машинных операций .

Представление чисел в ЭВМ

Целые числа, представимые в ЭВМ (целые машинные числа) - это все целые числа, принадлежащие замкнутому интервалу $[N^-, N^+]$.

Для целочисленного типа, занимающего **2 байта** (например, **integer** в языке Паскаль, - это $[-32768, 32767]$).

При вычислениях в ЭВМ чаще всего используется **представления числа в форме с плавающей запятой**, т.е. в виде

$$a = M \beta^p,$$

где **β** - основание системы счисления, **p** - порядок числа, **M** - мантисса числа .

Представление чисел в ЭВМ

Арифметические операции с целыми машинными числами выполняются на ЭВМ **точно**, если результат принадлежит интервалу $[N^-, N^+]$. В противном случае фиксируется ошибочная ситуация, а результат операции не определен или не имеет смысла. Поэтому при вычислениях следует заботиться о том, чтобы результаты арифметических операций не выходили за пределы интервала $[N^-, N^+]$.

Представление чисел в ЭВМ

Вещественные числа могут быть записаны в двух формах. Обычную форму записи числа в виде

$$x = \pm a_n a_{n-1} \dots a_1 a_0, a_{-1} a_{-2} \dots a_{-m}$$

называют записью *с фиксированной запятой*.

В позиционной системе счисления с основанием β эта запись означает, что

$$x = \pm a_n \cdot \beta^n + a_{n-1} \cdot \beta^{n-1} + \dots + a_1 \cdot \beta^1 + a_0 + a_{-1} \cdot \beta^{-1} + a_{-m} \cdot \beta^{-m}$$

Каждое из чисел a_k , называемых разрядами числа, может принимать одно из значений $\{0, 1, \dots, \beta-1\}$.

Представление чисел в ЭВМ

О числах, записанных в виде

$0,63750 * 10^6; 637,50 * 10^3; 6,3750 * 10^5; 6375,0 * 10^1$

говорят, что они записаны в *форме с плавающей запятой*.

Запись числа с плавающей запятой, как следует из примера, не является однозначной. Для устранения этой неоднозначности принято первый множитель брать меньше единицы, и он должен состоять только из значащих цифр.

Такая форма записи числа называется *нормализованной*.

Представление чисел в ЭВМ

Такая форма записи числа называется *нормализованной* -

$$a = M\beta^p$$

где β - основание системы счисления, p - целое число (положительное, отрицательное или нуль), называемое *порядком* числа, M - число с фиксированной запятой, называемое *мантиссой* числа. Первая после запятой цифра числа M всегда отлична от нуля: $(\beta^{-1} \leq |M| < 1)$

Так, в рассмотренном примере, *нормализованной формой записи* числа 63750 будет $0,63750 \cdot 10^6$. Заметим, что в этой записи все цифры мантиссы верные. Для числа - 0,00384 нормализованной будет форма - $0,384 \cdot 10^{-2}$.

Представление чисел в ЭВМ

Для записи **вещественных чисел** в ЭВМ отводится фиксированное число разрядов (разрядная сетка), в которой выделены разряды для **записи мантиссы, порядка, знаков мантиссы и порядка**. Таким образом, множество машинных вещественных чисел характеризуется следующими четырьмя целыми константами: основанием счисления **β , $\beta \geq 2$** , точностью **t** , нижней границей экспоненты **e_-** и верхней границей экспоненты **e_+** . Ненулевые машинные вещественные числа имеют вид:

$$a = \pm \left(\frac{d_1}{\beta} + \frac{d_2}{\beta^2} + \dots + \frac{d_t}{\beta^t} \right) \beta^p = \pm d_1 d_2 \dots d_t \cdot \beta^p$$

где

$$0 \leq d_i < \beta, d_1 \neq 0, e_- \leq p \leq e_+$$

Представление чисел в ЭВМ

Вещественный **машинный нуль** представляется в ЭВМ специальным образом. Обычно это число, у которого все цифры $d_1 = d_2 = \dots = d_t = 0$ или $p < e^-$.

Минимальное положительное число ε_0 , которое может быть представлено в ЭВМ, называется *машинным эпсилоном* ($\varepsilon_0 = \beta^{e^-}$).

Максимальное положительное число ε_∞ - *машинной бесконечностью* - $\varepsilon_\infty = \beta^{e^+} (1 - \beta^{-t})$.

Таким образом, машинные вещественные числа принадлежат замкнутому интервалу $-\varepsilon_\infty \leq x \leq \varepsilon_\infty$

На интервале $-\varepsilon_0 \leq x \leq \varepsilon_0$ лежит только одно машинное вещественное число – машинный нуль.

Представление чисел в ЭВМ

В различных языках программирования реализованы несколько типов вещественных чисел с различной длиной мантиссы. Например, в языке Паскаль реализованы типы *real*, *single*, *double*, *extended*, некоторые характеристики которых приведены ниже

Тип	Наименование	Занимаемая память	Число значащих цифр	Диапазон допустимых значений
<i>real</i>	вещественный	6 байт	11-12	$2.9 \cdot 10^{-39} \div 1.7 \cdot 10^{38}$
<i>single</i>	с одинарной точностью	4 байта	7-8	$1.5 \cdot 10^{-38} \div 3.4 \cdot 10^{38}$
<i>double</i>	с двойной точностью	8 байт	15-16	$5.0 \cdot 10^{-324} \div 1.7 \cdot 10^{308}$
<i>extended</i>	с повышенной точностью	10 байт	19-20	$3.4 \cdot 10^{-4932} \div 1.1 \cdot 10^{4932}$

Представление чисел в ЭВМ

Тип	Размер в байтах (битах)	Интервал изменения
char	1 (8)	от -128 до 127
unsigned char	1 (8)	от 0 до 255
signed char	1 (8)	от -128 до 127
int	2 (16)	от -32768 до 32767
unsigned int	2 (16)	от 0 до 65535
signed int	2 (16)	от -32768 до 32767
short int	2 (16)	от -32768 до 32767
unsigned short int	2 (16)	от 0 до 65535
signed short int	2 (16)	от -32768 до 32767
long int	4 (32)	от -2147483648 до 2147483647
unsigned long int	4 (32)	от 0 до 4294967295
signed long int	4 (32)	от -2147483648 до 2147483647
float	4 (32)	от 3.4E-38 до 3.4E+38
double	8 (64)	от 1.7E-308 до 1.7E+308
long double	10 (80)	от 3.4E-4932 до 3.4E+4932

Представление чисел в ЭВМ

Из-за конечности разрядной сетки в ЭВМ можно представить точно лишь конечное подмножество вещественных чисел. Рассмотрим два машинных вещественных числа f_1 и f_2 ($f_1 < f_2$) таких, что между ними нет других машинных чисел.

Справедливо неравенство

$$\frac{f_2 - f_1}{f_1} \leq \varepsilon_1 = \beta^{1-t}$$

Число $\varepsilon_1 = \beta^{1-t}$ зависит от разрядности мантииссы и характеризует **точность представления чисел в ЭВМ**.

Представление чисел в ЭВМ

Таким образом, любое вещественное число x , не представимое точно, заменяется ближайшим к нему числом \hat{x} , представимым точно. Такая замена называется *округлением числа x* . Величина погрешности замены не превышает $\varepsilon_1 x$ и зависит от способа округления.

Простейшим является *отбрасывание* всех разрядов мантиссы числа x , которые выходят за пределы разрядной сетки. Предельная относительная погрешность округления при таком способе равна

$$\delta(\hat{x}) \leq \varepsilon$$

Представление чисел в ЭВМ

При более точном *симметричном округлении* к последнему сохраняемому разряду мантиссы прибавляется единица, если первый отбрасываемый двоичный разряд равен 1. Предельная относительная погрешность округления в данном случае вдвое меньше:

$$\delta(\hat{x}) \leq \frac{1}{2} \beta^{-t+1} = \frac{1}{2} \varepsilon$$

Таким образом, можно считать, что машинные константы

$\varepsilon_0, \varepsilon_\infty, \varepsilon_1$ характеризуют погрешность представления вещественных чисел в ЭВМ.

Накопление погрешностей округления

Ошибочно полагать, что в силу малости величин ошибки округления практически не влияют на результаты вычислений. В процессе проведения вычислений погрешности округления могут накапливаться, так как выполнение каждой из четырех арифметических операций вносит некоторую погрешность.

Поскольку ЭВМ выполняет промежуточные вычисления с двойной точностью, т.е. с мантиссой, содержащей $2t$ разрядов, то округлению до t разрядов подвергается лишь окончательный результат. Таким образом, выполнение каждой арифметической операции вносит относительную погрешность, не превышающую $\varepsilon 1$.

ПРИМЕР. Рассмотрим процесс вычисления значений функции $\sin(t)$ с использованием частичных сумм ряда Тейлора

```
Series[Sin[x], {x, 0, 17}]
```

разлож... | синус

$$x - \frac{x^3}{6} + \frac{x^5}{120} - \frac{x^7}{5040} + \frac{x^9}{362880} - \frac{x^{11}}{39916800} + \frac{x^{13}}{6227020800} - \frac{x^{15}}{1307674368000} + \frac{x^{17}}{355687428096000} + O[x]^{18}$$

Получим это же разложение, используя оператор цикла:

```
YDT = {}; ydata[1] = t; n = 9;
```

```
For [i = 1, i ≤ n, i++,  
Цикл Для
```

```
ydata[i + 1] = ydata[i] +  $\frac{(-1)^i t^{2xi+1}}{(2 \times i + 1)!}$ ; YDT = Append[YDT, ydata[i]]; ];
```

Добавить в конец

матричная форма

$$\begin{array}{c}
 t \\
 t - \frac{t^3}{6} \\
 t - \frac{t^3}{6} + \frac{t^5}{120} \\
 t - \frac{t^3}{6} + \frac{t^5}{120} - \frac{t^7}{5040} \\
 t - \frac{t^3}{6} + \frac{t^5}{120} - \frac{t^7}{5040} + \frac{t^9}{362880} \\
 t - \frac{t^3}{6} + \frac{t^5}{120} - \frac{t^7}{5040} + \frac{t^9}{362880} - \frac{t^{11}}{39916800} \\
 t - \frac{t^3}{6} + \frac{t^5}{120} - \frac{t^7}{5040} + \frac{t^9}{362880} - \frac{t^{11}}{39916800} + \frac{t^{13}}{6227020800} \\
 t - \frac{t^3}{6} + \frac{t^5}{120} - \frac{t^7}{5040} + \frac{t^9}{362880} - \frac{t^{11}}{39916800} + \frac{t^{13}}{6227020800} - \frac{t^{15}}{1307674368000} \\
 t - \frac{t^3}{6} + \frac{t^5}{120} - \frac{t^7}{5040} + \frac{t^9}{362880} - \frac{t^{11}}{39916800} + \frac{t^{13}}{6227020800} - \frac{t^{15}}{1307674368000} + \frac{t^{17}}{355687428096000}
 \end{array}$$

Составим программу вычисления значения функции $y = \sin(t)$ с точностью ε . Используем тот факт, что **частичная сумма знакопередающегося ряда** отличается от точного значения функции не более чем **на величину первого отброшенного члена ряда**, т.е. прекратим вычисление суммы когда выполнится неравенство

$$\left| \frac{t^{2k+1}}{(2k+1)!} \right| \leq \varepsilon.$$

Вычисление очередного слагаемого организуем по рекуррентной формуле

$$a_{n+1} = a_n \cdot (-1) \cdot t^2 / (2n) / (2n+1)$$


```
TableForm[N[Table[{t, Sin[t]}, {t, 0, 1, 0.1}]]]
```


[табличная ... [... [таблица зн... [синус

```
TableForm[N[Table[{t, ydata[n]}, {t, 0, 1, 0.1}]]]
```

[табличная ... [... [таблица значений

0.	0.
0.1	0.0998334
0.2	0.198669
0.3	0.29552
0.4	0.389418
0.5	0.479426
0.6	0.564642
0.7	0.644218
0.8	0.717356
0.9	0.783327
1.	0.841471

0.	0.
0.1	0.0998334
0.2	0.198669
0.3	0.29552
0.4	0.389418
0.5	0.479426
0.6	0.564642
0.7	0.644218
0.8	0.717356
0.9	0.783327
1.	0.841471



0.	0.
1.	0.841471
2.	0.909297
3.	0.14112
4.	-0.756802
5.	-0.958924
6.	-0.279415
7.	0.656987
8.	0.989358
9.	0.412118
10.	-0.544021
11.	-0.99999
12.	-0.536573
13.	0.420167
14.	0.990607
15.	0.650288

0.	0.
1.	0.841471
2.	0.909297
3.	0.14112
4.	-0.7568
5.	-0.958776
6.	-0.274807
7.	0.740738
8.	2.01415
9.	9.67673
10.	65.3945
11.	385.619
12.	1930.77
13.	8433.03
14.	32 832.
15.	115 803.

Sin[0.5]

|синус

0.479426

MatrixForm[N[SDT]]

|матричная ф... |численное

MatrixForm=

$$\begin{pmatrix} 0.5 \\ 0.479167 \\ 0.479427 \\ 0.479426 \\ 0.479426 \\ 0.479426 \\ 0.479426 \end{pmatrix}$$

MatrixForm[N[YDT]]

|матричная ф... |численное

MatrixForm=

$$\begin{pmatrix} 0.5 \\ -0.0208333 \\ 0.000260417 \\ -1.5501 \times 10^{-6} \\ 5.38229 \times 10^{-9} \\ -1.22325 \times 10^{-11} \\ 1.96033 \times 10^{-14} \end{pmatrix}$$

N[Sin[15]]

... | синус

0.650288

N[YDT]

| численное приближение

{15., -562.5, 6328.13, -33900.7, 105940., -216695., 312540., -334865., 277002., -182238.,
97627.6, -43411.5, 16279.3, -5217.73, 1445.8, -349.79, 74.5291, -14.0916, 2.38034, -0.361388,
0.0495807, -0.006177, 0.000701932, -0.0000730502, 6.98822×10^{-6} , -6.16608×10^{-7} , 5.03399×10^{-8} ,
 -3.81363×10^{-9} , 2.68818×10^{-10} , -1.76751×10^{-11} , 1.08658×10^{-12} , -6.25911×10^{-14} , 3.38533×10^{-15} }

N[SDT]

| численное приближение

{15., -547.5, 5780.63, -28120., 77819.5, -138875., 173665., -161199., 115803.,
-66435.5, 31192., -12219.4, 4059.87, -1157.86, 287.94, -61.8502, 12.6789,
-1.41271, 0.967637, 0.606249, 0.65583, 0.649653, 0.650354, 0.650281, 0.650288,
0.650288, 0.650288, 0.650288, 0.650288, 0.650288, 0.650288, 0.650288, 0.650288}

N[YDT]

[численное приближение]

{45., -15187.5, 1.53773×10^6 , -7.41408×10^7 , 2.08521×10^9 , -3.83868×10^{10} , 4.9829×10^{11} ,
 -4.80494×10^{12} , 3.57721×10^{13} , -2.11808×10^{14} , 1.02122×10^{15} , -4.08689×10^{15} , 1.37933×10^{16} ,
 -3.97883×10^{16} , 9.92257×10^{16} , -2.16056×10^{17} , 4.14312×10^{17} , -7.05026×10^{17} , 1.07183×10^{18} ,
 -1.46455×10^{18} , 1.80836×10^{18} , -2.02764×10^{18} , 2.07373×10^{18} , -1.94232×10^{18} , 1.67228×10^{18} ,
 -1.32798×10^{18} , 9.75751×10^{17} , -6.65285×10^{17} , 4.22056×10^{17} , -2.49755×10^{17} , 1.38184×10^{17} ,
 -7.16393×10^{16} , 3.48725×10^{16} , -1.59694×10^{16} , 6.89218×10^{15} , -2.80818×10^{15} , 1.08192×10^{15} ,
 -3.94754×10^{14} , 1.36599×10^{14} , -4.48901×10^{13} , 1.40282×10^{13} , -4.17382×10^{12} , 1.18375×10^{12} ,
 -3.20382×10^{11} , 8.28363×10^{10} , -2.04815×10^{10} , 4.84748×10^9 , -1.09923×10^9 , 2.3904×10^8 ,
 -4.98925×10^7 , 1.00032×10^7 , -1.92809×10^6 , 357 543., -63 835.8, 10 980.9, -1821.16,
291.392, -45.009, 6.71553, -0.968449, 0.135063, -0.0182262, 0.00238116, -0.000301328,
0.0000369543, -4.39415×10^{-6} , 5.06844×10^{-7} , -5.67363×10^{-8} , 6.16632×10^{-9} , -6.50965×10^{-10} ,
 6.67783×10^{-11} , -6.65941×10^{-12} , 6.45848×10^{-13} , -6.09376×10^{-14} , 5.5958×10^{-15} }