

Основы теории погрешности вычислений

Лекция 2



План

- Абсолютная и относительная погрешности
- Запись и округление чисел
- **Способы оценки погрешностей вычислений**
 - Метод границ
 - Дифференциальная оценка
- Обратная задача теории погрешностей
- Инструменты приближенных вычислений

Абсолютная погрешность

- Пусть X – точное значение некоторой величины (обычно неизвестное), x – число, принятое за ее приближенное значение ($X \approx x$).
- Разность $X - x$ называют погрешностью приближенного значения x , а модуль

$$\varepsilon_x = |X - x|$$

- абсолютной погрешностью величины x .

Предельные погрешности

Предельной *абсолютной погрешностью*, являющейся верхней границей, называют такое по возможности наименьшее число, для которого справедливо неравенство:

$$\Delta x \leq |X - x|$$

Пример.

Число $\pi = 3.141592653589\dots$

может округляться по-разному в зависимости от потребностей задачи:

$$\pi \approx 3.14 \quad \Delta\pi = 0.0016 \text{ или } \Delta\pi = 0.002$$

$$\pi \approx 3.1416 \quad \Delta\pi = 0.000008 \text{ или } \Delta\pi = 0.00001$$

Предельные погрешности

Предельная *относительная погрешность* приближенного числа – это отношение предельной абсолютной погрешности к абсолютному значению приближения:

$$\delta x \geq \frac{\Delta x}{|x|}$$

Относительную погрешность обычно выражают в процентах.

Пример.

Относительная погрешность округления
числа $\pi \approx 3.14$:

$$\delta\pi \geq \frac{0.0016}{3.14}, \delta\pi \geq 0.0005095\dots$$

В процентах:

$$\delta\pi \geq 0.05095\dots\%$$

*Округление погрешностей рассмотрим
позже.*

Границы точного результата

Если известны x и Δ_x , то принято записывать:

$$X \in [x - \Delta_x, x + \Delta_x]$$

или

$$X = x \pm \Delta_x$$

Опр. Любая пара чисел $НГ_x$ и $ВГ_x$, такая, что $НГ_x \leq x \leq ВГ_x$ называется **нижней и верхней границей** приближенного числа x .

Пример.

Ширина шкафа S известна
приблизенно с погрешностью Δ_s ,
ширина двери H и её погрешность Δ_h .
При каком условии шкаф
гарантированно пройдет в дверь?

$$ВГ_{\text{шкафа}} < НГ_{\text{двери}}$$

$$s + \Delta s < h - \Delta h$$



Значащие цифры

Определение. Все цифры десятичной записи числа, начиная с первой ненулевой слева, называются **значащими**.

- ***308,6170 – 7 значащих цифр***
 - ***0,00235 – 3 значащих цифры***
- !!! 57 000 – 5 з.ц., $5,7 \cdot 10^4$ – 2 з.ц., $570 \cdot 10^2$ – 3 з.ц.***



Упражнение

Определите количество значащих цифр в записи чисел:

1. 0,0050

2

2. - 38,412

5

3. 4100

4

4. $20 \cdot 10^5$

2

Округление чисел

Определение: *округлением* числа называют замену его близким по величине, но с меньшим количеством значащих цифр.

Различают 3 вида округления:

Симметричное округление – к ближайшему числу: $3,6 \approx 4$; $3,67 \approx 3,7$; $3,4 \approx 3$; $3,42 \approx 3,4$

С избытком - к большему числу: $3,6 \approx 4$; $3,2 \approx 4$

С недостатком - к меньшему числу
(округление отсечением): $3,67 \approx 3,6$; $3,2 \approx 3$.

Округление погрешностей

Правило:

В записи погрешности обычно оставляют *только 1-2 значащие цифры.*

Для сохранения условия, соответствующего определению предельной погрешности, округление её всегда производят **с избытком.**

Это правило распространяется как на абсолютную, так и на относительную погрешности.

Пример.

В рассмотренном ранее примере
вычисления площади относительная
погрешность должна была округляться
следующим образом:

$$\delta\pi \geq 0.05095\dots\%$$

$$\delta\pi = 0.051\% \quad 0.068\% =$$

НГ приближенного числа округляют с недостатком, ВГ - с избытком.

Абсолютная погрешность приближенного числа, полученного при округлении с недостатком (избытком) равна единице последнего сохраненного разряда:

$$1,2776 \approx 1,27; \Delta = 0,01.$$

Симметричное округление дает меньшую ошибку - половину единицы последнего разряда: $1,2776 \approx 1,28; \Delta = 0,005.$

Верные цифры

Определение. Цифра в записи числа a , называется *верной*, если погрешность не превосходит единицы её разряда; цифра *верная в строгом смысле* - если погрешность не превосходит половины её разряда.

Пример

Определить верные цифры:

$$a = 18.572, \Delta_a = 0.08$$

Решение:

1: $10 \geq 0.08$ - верная

8: $1 \geq 0.08$ - верная

5: $0.1 \geq 0.08$ - верная

7: $0.01 \leq 0.08$ - сомнительная

2: $0.001 \leq 0.08$ - сомнительная

} в широком смысле

Пример

Для приближённого числа $x=72,256$ известна абсолютная погрешность

$$\Delta_x = 0,01.$$

Требуется определить его верные значащие цифры в а) широком и б) строгом смыслах.

$$x=72,256$$

$$\Delta_x = 0,01.$$

а) Решение:

2: $0.1 \geq 0.01$ - верная

5: $0.01 = 0.01$ – верная

6: $0.001 \leq 0.01$ - сомн.

Ответ: верные в
широком смысле
цифры 7;2;2;5

а 6 - сомнительная

б) Решение:

2: $0.05 \geq 0.01$ - верная

5: $0.005 \leq 0.01$ – сомн.

Ответ: верные в
строгом смысле
цифры 7;2;2

а 5 и 6 - сомнительные

Верная цифра не обязательно буквально совпадает с соответствующей цифрой точного числа.

$X = 1,999$ – точное число, $x = 2,000$ – его приближение. Тогда $\Delta_x = 0,001$, и, следовательно, три первые цифры верные, хотя ни одна из них не совпадает с соответствующими цифрами точного числа.

Если исходные данные приводятся без указания погрешностей, но с известными верными цифрами, то погрешность можно определить, исходя из определения верной цифры.

Пример. $X=4,06$ $\Delta_x=0,01$ (в широком см.)

$X=4,06$ $\Delta_x=0,005$ (строгом см.)

Если не уточняется трактовка смысла (широкая, строгая), то по умолчанию принимается ***строгий*** вариант.

Запись приближенных чисел

Точность приближенного числа зависит не от количества значащих цифр, а от количества верных значащих цифр.

Если полученный в результате вычислений результат содержит излишнее количество сомнительных значащих цифр, то его округляют.

Запись приближённых чисел

Правило: в промежуточных результатах вычислений обычно сохраняют 1-2 сомнительные цифры, а окончательные результаты округляют с сохранением не более одной сомнительной цифры.



Упражнение

Приближённое значение $x=24,6035$ имеет относительную погрешность $\delta=0,1\%$. Найти Δ_x и округлить число x с точностью до верных цифр.

Решение

$$\delta_x = \frac{\Delta_x}{|x|} \cdot 100\%$$

$$\Delta_x = \frac{\delta_x}{100\%} \cdot |x|$$

$$\Delta_x = \frac{0,1}{100\%} \cdot 24,6035 = 0,0246035 \approx 0,03$$

$$6 : 0,05 \geq 0,03$$

$$0 : 0,005 \leq 0,03$$

Верные цифры, округляем

24,6 ≈



Оценка погрешности вычисления функции

Определим, как вычислить погрешность функции, некоторые аргументы которой заданы приближенно.

Задачу нахождения погрешности функции по заданным погрешностям приближенных аргументов называют **основной задачей** теории погрешностей.

Оценка погрешности по способу границ

Пусть $y=f(x)$ - функция, для которой необходимо найти погрешность, a - приближенное исходное данное и известны $НГ_a$ и $ВГ_a$.

Необходимо определить $y=f(x)$, $НГ_y$ и $ВГ_y$.

Для нахождения границ результата вычисляют $y_1=f(НГ_x)$ и $y_2=f(ВГ_x)$, а затем меньшее из этих значений принимают за $НГ_y$, а большее за $ВГ_y$ и округляют нижнее с недостатком, а верхнее с избытком.

Пример:

$$y = \frac{e^x}{1+x^2}$$

$$1,25 < x < 1,28,$$

Решение.

$$НГ(x) = \frac{e^{1,25}}{1+1,25^2} = 1,3620849\dots$$

$$НГ_y = 1.362$$

$$ВГ(x) = \frac{e^{1,28}}{1+1,28^2} = 1,3631896\dots$$

$$ВГ_y = 1.364$$

$$\Delta_y = \frac{ВГ_y - НГ_y}{2} = \frac{1,364 - 1,362}{2} = 0,001$$

$$y \approx \frac{ВГ_y + НГ_y}{2} \approx 1,363$$

$$y = 1,363 \pm 0,001$$

Рассмотрим случай двух переменных

$$H\Gamma_a \leq a \leq B\Gamma_a$$

$$z = f(a, b) - ?$$

$$H\Gamma_b \leq b \leq B\Gamma_b$$

$$z = f(H\Gamma_a, H\Gamma_b)$$

$$z = f(H\Gamma_a, B\Gamma_b)$$

$$z = f(B\Gamma_a, B\Gamma_b)$$

$$z = f(B\Gamma_a, H\Gamma_b)$$

} min \longrightarrow $H\Gamma_z$, max \longrightarrow $B\Gamma_z$

Замечание

Прежде чем производить расчет для всех возможных вариантов (для n – переменных это 2^n), необходимо попытаться оценить характер зависимости от некоторых переменных.

Пример. Как оценить, какие границы аргументов использовать для $НГ_F$ и $ВГ_F$?

$$f(a, b, c) = \frac{a^2 + \sqrt{b}}{c + \frac{e^a}{b}}$$



Дифференциальная оценка погрешности

Теорема. Пусть x, y являются приближениями к точным значениям аргументов X, Y с абсолютными погрешностями Δ_x и Δ_y . Если функция $z=f(x,y)$ дифференцируема, M_x и M_y – максимумы частных производных

$$\left| \frac{\partial f(x, y)}{\partial x} \right| \text{ и } \left| \frac{\partial f(x, y)}{\partial y} \right|$$

в прямоугольнике $\begin{cases} x - \Delta x \leq x \leq x + \Delta x \\ y - \Delta y \leq y \leq y + \Delta y \end{cases}$

то абсолютная *дифференциальная погрешность функции*

$$\Delta_z \leq M_1 \cdot \Delta_x + M_2 \cdot \Delta_y$$

Дифференциальная оценка погрешности

Замечание. При малых значениях Δ_x и Δ_y можно вместо максимумов частных производных в прямоугольнике брать значения производных с приближенными значениями аргументов:

$$\Delta_z = |f'_x(x, y)| \cdot \Delta_x + |f'_y(x, y)| \cdot \Delta_y.$$

Пример

Пусть $x = -0,68 \pm 0,004$, $y = 1,134 \pm 0,0003$.

Требуется найти значение z для $f(x, y) = x^2 + \sin y$ и оценить погрешность дифференциальным способом.

$$f(x,y) = x^2 + \sin y$$

$$\frac{\partial f}{\partial x} = 2 \cdot x,$$

$$\frac{\partial f}{\partial y} = \cos(y).$$

$$\Delta_z = |2 \cdot (-0,68)| \cdot 0,004 + |\cos(1,134)| \cdot 0,003 = 0,00556691163\dots$$

$$\approx 0,006$$

$$z = (-0,68)^2 + \sin(1,134) = 1,368511588\dots$$

$$\approx 1,37$$

$$z = 1,37 \pm 0,006.$$

$$\delta = \frac{0,006}{1,37} = 0,004379562\dots$$

$$\approx 0,0044 = 0,44\%$$

Пример

Пусть в выражении

$$d = 2,63 - 1,026 \cdot \sqrt{5,40}$$

все числа приближённые и записаны верными цифрами. Требуется найти значение d и определить абсолютную и относительную погрешности.

Решение

Для функции трех переменных:

$$d = 2,63 - 1,026 \cdot \sqrt{5,40}$$

$$f(x, y, z) = x - y \cdot \sqrt{z}$$

$$f(x, y, z) = x - y \cdot \sqrt{z}$$

$$f'_x = 1$$

$$f'_y = -\sqrt{z}$$

$$f'_z = -\frac{y}{2 \cdot \sqrt{z}}$$

$$\Delta_d = |1| \cdot 0,01 + |-\sqrt{5,40}| \cdot 0,001 + \left| -\frac{1,026}{2 \cdot \sqrt{5,40}} \right| \cdot 0,01 =$$

$$= 0,01 + 0,002323790007... + 0,0022076... \approx 0,0145...$$

$$\Delta_d = 0,015, \quad d = 2,63 - 1,026 \cdot \sqrt{5,40} = 0,24579146 \approx 0,25.$$

$$\delta_d = \frac{0,015}{0,25} = 0,06 = 6\%.$$



Обратная задача теории погрешностей

Обратная задача теории погрешностей состоит в том, что по заданной абсолютной погрешности функции необходимо определить, каковы должны быть абсолютные погрешности ее аргументов.

Обратная задача теории погрешностей

Одна и та же суммарная оценка погрешности функции нескольких аргументов, вычисляемая, например, дифференциальным способом, может быть получена при различных распределениях погрешностей аргументов.

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

Обратная задача теории погрешностей

Для решения обратной задачи обычно пользуются принципом «равных влияний»:

$$\frac{\Delta f}{n} = \left| \frac{\partial f}{\partial x_i} \right| \cdot \Delta x_i, \text{ т.е. все слагаемые оценки равны.}$$

Тогда

$$\Delta x_i = \frac{\Delta f}{n} / \left| \frac{\partial f}{\partial x_i} \right|.$$



Инструменты приближенных вычислений

Электронные таблицы

Математические пакеты

Среды программирования

Точность представления чисел в компьютере

Для вещественных чисел используется нормализованное представление в форме с плавающей запятой

$$x = \pm M \cdot 10^P$$

мантисса

порядок

Точность представления чисел зависит от длины разрядной сетки. Данные с одинарной точностью имеют не более 7 верных десятичных знаков, с двойной точностью – 16.

Вычисление функций



```
1 import math
2
3 def f(x, y):
4     return x**2 + math.sin(y)
5
6 print(f(-0.68, math.pi/3))
```

```
1 from math import sin, pi
2
3 def f(x, y):
4     return x**2 + sin(y)
5
6 print(f(-0.68, pi/3))
```

```
>>> %Run L1_ex_2.py
1.3284254037844387
```



$f := (x, y) \rightarrow x^2 + \sin(y)$
 $(x, y) \rightarrow x^2 + \sin(y)$
 $f\left(-0.68, \frac{\text{Pi}}{3}\right)$
 $0.4624 + \frac{1}{2} \sqrt{3}$
evalf(%)
1.328425404

Вычисление производных



```
1 """
2 Производная функции в символьном представлении
3
4 """
5
6 import sympy as sym
7
8 x = sym.Symbol('x')
9
10 f = sym.exp(x) / (1 + x**2)
11 df = f.diff(x)
12
13 sym.pprint(f)
14 sym.pprint(df)
15
```

Результат:

$$\frac{e^x}{x^2 + 1}$$
$$-\frac{2 \cdot x \cdot e^x}{(x^2 + 1)^2} + \frac{e^x}{x^2 + 1}$$

Вычисление производных



```
1 """
2 Вычисление функции и производной
3
4 """
5
6 import sympy as sym
7
8 x = sym.Symbol('x')
9
10 f = sym.exp(x) / (1 + x**2)
11 df = f.diff(x)
12
13 print('Функция: ', f.subs(x, 2.4), 'Производная: ', df.subs(x, 2.4))
14
```

Результат:

Функция: 1.63064739358604 Производная: 0.472791256128495



Функции округления



- **round(number [, ndigits])** - округляет число number до ndigits знаков после запятой (по умолчанию, до нуля знаков, то есть, до ближайшего целого)
- **math.ceil(number)** — округление до ближайшего большего числа
- **math.floor(number)** - округление вниз
- **math.trunc(number)** - отсекает значение X до целого
- **round(x)** - округляет x до ближайшего целого
- **ceil(x)**
- **floor(x)**
- **trunc(x)** - возвращает:
floor(x) для $x \geq 0$
floor(-x) для $x < 0$

Примеры

```
>>> round(-11.416051)
-11

>>> round(2.75645)
3

>>> round(3.1415, 2)
3.14
```

```
>>> 0.2 + 0.3 == 0.5
True

>>> 0.2 + 0.1 == 0.3
False
```

```
>>> 0.2 + 0.1
0.30000000000000004
```

```
>>> round(3.5)
4

>>> round(8.5)
8
```

округление происходит до
ближайшего чётного
(банковское округление)

Особенности машинной арифметики

В машинных вычислениях:

- числа представлены с ограниченным числом разрядов
- выполняется огромное число арифметических операций, что приводит к накоплению ошибок
- промежуточные результаты обычно не отражаются.

Источники вычислительной погрешности

1. Представление чисел в 2-й системе (конечное 10-ое число может стать бесконечным.)
2. Результаты отдельных арифметических операций не подвергаются правильному округлению (для сокращения задержек из-за переноса единицы в старшие разряды).
3. Ограниченный диапазон чисел, представимых в разрядной сетке компьютера (проблема машинного нуля и переполнения).

Рекомендации

Для исключения неблагоприятного влияния на порядок организации вычислений прибегают к следующим приемам, уменьшающим вычислительную погрешность:

- суммирование нужно начинать с **малых по модулю слагаемых**: в противном случае они могут оказаться несоизмеримыми с накопленной суммой и не окажут на нее должного влияния
- следует избегать **вычитания близких чисел**, при котором происходит катастрофическая потеря верных цифр.
- последовательное умножение упорядоченных чисел может также привести к потере точности, поэтому целесообразно нарушать **порядок** умножения таких сомножителей
- нельзя **сравнивать на равенство** числа в нормализованном формате

