Тема 2. ЭЛЕМЕНТЫ ТЕОРИИ ПОГРЕШНОСТЕЙ.

ОСНОВНЫЕ ПОНЯТИЯ И ОПРЕДЕЛЕНИЯ.

Применение любого численного метода неминуемо приводит к погрешности результатов решения задачи. Выделяют три основных составляющих погрешности при численном решении исходной задачи:

- ***неустранимая погрешность, связанная с неточным заданием исходных данных (начальные и граничные условия, коэффициенты и правые части уравнений);***
- ***погрешность метода, связанная с переходом к дискретному аналогу исходной задачи;***
- ***ошибка округления, связанная с конечной разрядностью чисел, представляемых в ЭВМ.***

Численный, или приближенный, метод реализуется всегда в виде вычислительного алгоритма. Прежде всего, алгоритм должен быть реализуем - обеспечивать решение задачи за допустимое машинное время. Важной характеристикой алгоритма является его погрешность. Для очень малых значений погрешности время вычислений может быть недопустимо большим. Поэтому на практике добиваются некоторого компромисса между точностью и затрачиваемым машинным временем.

Если погрешность в процессе вычислений неограниченно возрастает, то такой алгоритм называется неустойчивым, или расходящимся. В противном случае алгоритм называется устойчивым, или сходящимся.

ОСНОВНЫЕ ПОНЯТИЯ И ОПРЕДЕЛЕНИЯ.

Существуют четыре источника погрешности результата:

1. ***Погрешность математической модели***

(возникает из-за стремления обеспечить сравнительную простоту её технической реализации и доступности исследования)

2. ***Погрешность исходных данных*** (неустранимая с точки зрения вычислительного эксперимента, погрешность, но эту погрешность возможно оценить для выбора алгоритма расчета и точности вычислений. Как известно, ошибки эксперимента условно делят на систематические, случайные и грубые, а идентификация таких ошибок возможна при статистическом анализа результатов эксперимента**).**

3. ***Погрешность численного метода***

(связана, например, с заменой интеграла суммой, с усечением рядов при вычислении функций, с интерполированием табличных значений функциональных зависимостей и т. п. Как правило, погрешность численного метода регулируема и может быть уменьшена до любого **разумного** значения путём изменения алгоритма вычислений или параметра)

4. ***Вычислительная погрешность (погрешность округления)*** (возникает из-за округления чисел, промежуточных и окончательных результатов счета. Она зависит от правил и необходимости округления, а также от алгоритмов численного решения).

ОСНОВНЫЕ ПОНЯТИЯ И ОПРЕДЕЛЕНИЯ.

Пусть имеется реальный маятник, совершающий затухающие колебания, начинающий движение в момент $t = t_0$. Требуется найти угол отклонения φ от вертикали в момент t_1. Движение маятника описывается дифференциальным уравнением:

$$\boxed{l\frac{d^2\varphi}{dt^2} + g \cdot \sin(\varphi) + \mu\frac{d\varphi}{dt} = 0}$$

где l – длина маятника, g – ускорение силы тяжести, μ – коэффициент трения.

Как только принимается такое описание задачи, решение уже приобретает неустранимую погрешность, в частности потому, что реальное трение зависит от скорости не совсем линейно (**погрешность модели**). Кроме того, воспроизведя реальный эксперимент, мы зададим g (в известной точке планеты), μ с некоторой точностью, и получим набор значений с погрешностью, которую можем оценить из анализа статистики некоторого числа однотипных опытов (**погрешность исходных данных**).

Взятое в модели дифференциальное уравнение нельзя решить в явном виде, для его решения требуется применить какой-либо численный метод, имеющий заранее известную погрешность.

После совершения вычислений мы получим значения с погрешностью большей, нежели погрешность метода, так как к ней прибавится **погрешность округления**.

ОСНОВНЫЕ ПОНЯТИЯ И ОПРЕДЕЛЕНИЯ.

Абсолютная погрешность некоторого числа a равна разности между его истинным и приближенным значением a^*, полученным в результате вычисления или измерения. $\Delta(a^*) = a - a^*$

Относительная погрешность – это отношение абсолютной погрешности к приближенному значению числа $\delta(a^*) = \Delta a^* / |a|$

Относительную погрешность обычно выражают в процентах: $\boxed{\delta(a^*) = \dfrac{\Delta a^*}{|a|} \cdot 100\%}$

Использование относительных погрешностей удобно тем, что они не зависят от масштабов величин и единиц измерения.

Пусть известно $\Delta a^* = 0{,}1$- это большая или малая погрешность?

Если $a^* \approx 0{,}33$, то скорее всего погрешность велика, если $a^* \approx 0{,}33 \cdot 10^6$ следует признать её малой.

При оперировании относительной величиной погрешности имеем:

$\delta(0{,}33) = 33\%$ в первом случае и $\delta(0{,}33 \cdot 10^6) = 0{,}33 \cdot 10^{-5}\%$ во втором

Таким образом, для оценки погрешностей логичней пользоваться её относительной мерой.

ОСНОВНЫЕ ПОНЯТИЯ И ОПРЕДЕЛЕНИЯ.

В практике решения задач часто используется термин ***«точность».***

Точность в качественных рассуждениях подразумевается как противоположность ***погрешности***, хотя для их количественных измерений используются одни и те же характеристики.

Повышение ***точности*** величины означает уменьшение её ***погрешности .***

При численном решении любой практической задачи необходимо всегда указывать требуемую точность результата : «***требуется найти решение с заданной точностью*** ε ».

В связи с этим необходимо уметь:

1) зная заданную точность исходных данных, оценивать точность результата (прямая задача теории погрешностей);

2) зная требуемую точность результата, выбирать необходимую точность исходных данных (обратная задача теории погрешностей).

Приближенные числа и оценка их погрешностей

Истинное значение величины a обычно неизвестно. Поэтому приведённые выражения для определения погрешностей зачастую практически не могут быть использованы, поскольку известно лишь приближенное значение a^* и нужно найти его ***предельную погрешность*** Δa^*.

Если значение Δa^* принимается в качестве абсолютной погрешности приближенного числа a, то в этом случае считается что истинное значение a находится в интервале:

$$(a - \Delta a^*) \le a \le (a + \Delta a^*).$$

Для приближенного числа, полученного в результате округления, абсолютная погрешность Δa^* ***принимается равной половине единицы последнего разряда числа.***

Например: Число $a = 0.734$ могло быть получено округлением чисел 0.73441 или 0.73353, тогда его предельная погрешность будет равна $\Delta a^* = 0{,}0005$.

Если при вычислениях на компьютере числа не округляются, а цифры, выходящие за разрядную сетку машины, ***отбрасываются***, то максимально возможная погрешность результата выполнения операции принимается ***в два раза больше*** по сравнению со ***случаем округления***.

Если число $a = 0.734$ было получено отбрасыванием например из числа 0.73453, тогда его предельная погрешность будет равна $\Delta a^* = 0{,}001$.

Правила записи приближенных чисел

Запись приближенных чисел должна подчиняться правилам, связанным с понятиями ***верных*** и ***значащих цифр***.

Любое десятичное число a^* может быть представлено в виде

$$a^* = \pm\alpha_n 10^n + \alpha_{n-1}10^{n-1} + \ldots \alpha_1 10 + \alpha_0 + \alpha_{-1}10^{-1} + \alpha_{-2}10^{-2} + \ldots + \alpha_{-m}10^{-m},$$

где α_i – цифры числа, 10^i – их позиция $(\pm i)$.

Например:

$$1358{,}7604 = 1\cdot 10^3 + 3\cdot 10^2 + 5\cdot 10 + 8 + 7\cdot 10^{-1} + 6\cdot 10^{-2} + 0\cdot 10^{-3} + 4\cdot 10^{-4}$$

Первая слева отличная от нуля цифра числа a^* и все расположенные справа от неё цифры называются ***значащими***, т.е. числа $25{,}047$ и $(-0{,}00250)$ имеют соответственно 5 и 3 значащих цифры.

Последнее число может быть записано $(-2{,}50\cdot 10^{-3})$ (запись $-2{,}5\cdot 10^{-3}$ будет неправильной).

Значащая цифра α_i называется ***верной***, если абсолютная погрешность числа не превосходит $1/2$ единицы разряда, соответствующего этой цифре, т.е. $\Delta a^* \le 0{,}5\cdot 10^i$, где 10^i указывает номер разряда $(\pm i)$.

Правила записи приближенных чисел

Пусть $a^* = 12{,}396$ и известна его предельная погрешность $\Delta a^* = 0{,}03$

Согласно определению здесь: $\Delta a^* > 0{,}5 \cdot 10^{-2}$

Верными цифрами числа a^* будут (**1, 2, 3**) а **9** и **6** ***сомнительные***.

Пусть $a = 0{,}0\underline{37862}$ и $\Delta a^* = 0{,}07$. Здесь $\Delta a^* > 0{,}5 \cdot 10^{-1}$.

Значит все значащие цифры ($3{,}7{,}8{,}6{,}2$) ***сомнительные.***

Следовательно можно сделать вывод:

Если число записано с указанием его абсолютной погрешности, то число верных знаков можно отсчитывать от первой значащей цифры числа до первой значащей цифры его абсолютной погрешности.

$$S = \underline{20{,}7}428 \qquad \Delta S^* = 0{,}0926$$

Первая значащая цифра числа - 2 (***первая слева отличная от нуля***)

Первая значащая цифра его абсолютной погрешности - 9 (***третья***)

Верными цифрами числа S будут первые три цифры - $(2, 0, 7)$

Оперировании понятиями верных значащих цифр.

Существуют определённые соглашения при оперировании понятиями верных значащих цифр.

1. ***Если число имеет лишь верные цифры, то и его округление имеет также только верные цифры*.**

2. ***Совпадение приближенного значения, имеющего все верные значащие цифры, с точным значением не обязательно*.**

3. ***Абсолютные и относительные погрешности числа принято округлять в большую сторону, так как при округлениях границы неопределённости числа, как правило, увеличиваются*.**

4. ***При изменении формы записи числа количество значащих цифр не должно меняться, т.е. необходимо соблюдать равносильность преобразований, например***

$7500 = 0{,}7500 \cdot 10^4$ и $0{,}110 \cdot 10^2 = 11{,}0$ – равносильные преобразования;

$7500 = 0{,}75 \cdot 10^4$ и $0{,}110 \cdot 10^2 = 11$ – неравносильные преобразования.

5. ***При вычислениях желательно сохранять такое количество значащих цифр, чтобы их число не превышало числа верных цифр более чем на одну – две единицы.***

6. ***Верные значащие цифры числа характеризуют <u>ориентировочно</u>***

Приближенные числа и оценка их погрешностей

Примеры оценки абсолютной погрешности величины a:

a ...	51,7	–0,0031	16
Δa^*...	0,05	0,00005	0,5

Предельное значение относительной погрешности - отношение предельной абсолютной погреш-ности к абсолютной величине приближенного числа: $\Delta(a^*) = \Delta a^* / |a|$

Например: $\delta(-2{,}3) = {}^{0{,}05}/_{|-2{,}3|} \approx 0{,}022 = (2{,}2\%)$.

Приведённые оценки погрешностей приближенных чисел справедливы, если в записи этих чисел все значащие цифры верные.

Напомним, что ***значащими цифрами*** считаются все цифры данного числа, начиная с ***первой слева ненулевой цифры.***

Например:

В числе **0.037** две значащие цифры - **3** и **7**, а в числе **14.80** все четыре цифры значащие. Кроме того, при изменении формы записи числа (например, при записи в форме с плавающей точкой) число значащих цифр не должно меняться, т.е. нужно соблюдать равносильность преобразований.

Например, записи $7500 = 0.7500 \cdot 10^4$ и $0.110 \cdot 10^2 = 11.0$ равносильные

записи $7500 = 0.75 \cdot 10^4$ и $0.110 \cdot 10^2 = 11$ неравносильные.

Правила расчёта погрешности округления

1. ***При сложении или вычитании чисел складываются их абсолютные погрешности:*** $\Delta(a \pm b) = \Delta a + \Delta b$

2. ***При умножении или делении чисел друг на друга складываются их относитель-ные погрешности:*** $\delta(a \cdot b) = \delta a + \delta b$ $\quad \delta(a/b) = \delta a + \delta b$

3. ***При возведении в степень приближенного числа его относительная погрешность умножается на показатель степени:*** $\delta(a^k) = k\delta a$

4. ***Относительная погрешность суммы положительных слагаемых заключена между наибольшим и наименьшим значениями относительных погрешностей этих слагаемых***.

$a>0,\ b>0,\ m = \min(\delta a, \delta b),\ M = \max(\delta a, \delta b)$ тогда $\delta(a+b) = M$

Пример: Оценить относительную погрешность функции $\boxed{y = \sqrt{\dfrac{a+b}{x^3(1-x)}}}$

$$\delta y = \tfrac{1}{2}[\delta(a+b) + 3\delta x + \delta(1-x)] = \tfrac{1}{2}\left[\frac{\Delta a + \Delta b}{|a+b|} + 3\frac{\Delta x}{|x|} + \frac{\Delta x}{|1-x|}\right]$$

Полученная оценка относительной погрешности содержит в знаменателе выражение $|1-x|$.

Ясно, что при $x \approx 1$ можно получить очень большую погрешность в результате деления на очень малое число.

Правила расчёта погрешности округления

Вычисление погрешности функций

$$y = f(x) \qquad u = f(x,y,z)$$

Пусть a - приближенное значение аргумента x, Δx - его абсолютная погрешность.

Абсолютной погрешностью функции будем считать её приращение, которое она испытывает при изменении аргумента на Δx.

Это приращение можно заменить дифференциалом: $\Delta y \approx dy$. Тогда для оценки абсолютной погрешности получим выражение $\Delta y = |f'(x)|\Delta x$

Аналогичное выражение можно записать для функции нескольких аргументов.

$$\Delta u = |f'_x(x,y,z)|\cdot\Delta x + |f'_y(x,y,z)|\cdot\Delta y + |f'_z(x,y,z)|\cdot\Delta z$$

здесь $\Delta x, \Delta y, \Delta z$ - абсолютные погрешности аргументов.

$$\delta u = \frac{\Delta u}{|f(x,y,z|}$$

Уменьшение погрешностей

При любых расчётах надо устанавливать такую точность вычислений, чтобы погрешность округления была существенно меньше всех остальных погрешнос-тей.

В связи с этим рассмотрим подробнее случай вычитания близких чисел.

Запишем выражение для относительной погрешности разности двух чисел

$$\delta(a-b) = \frac{\Delta(a-b)}{|a-b|} = \frac{\Delta a - \Delta b}{|a-b|}$$

При $a \approx b$ эта погрешность может быть сколь угодно большой.

Пример: Пусть $a = 2520$, $b = 2518$. В этом случае имеем абсолютные погрешности исходных данных $\Delta a = \Delta b = 0{,}5$ и относительные погрешности $\delta a \approx \delta b \approx 0{,}5/2518 \approx 0{,}0002$ $(0{,}02\%)$.

Относительная погрешность разности $(a\text{-}b)$

$$\delta(a-b) = \frac{0.5+0.5}{2520-2518} = \frac{1}{2} = 0.5\ (50\%)$$

Вывод: ***При организации вычислительных алгоритмов следует избегать вычитания близких чисел. При возможности алгоритм нужно видоизменить во избежание потери точности на некотором этапе вычислений.***

Уменьшение погрешностей

Пусть требуется найти сумму пяти четырёхразрядных чисел:

$$S = 0{,}2764 + 0{,}3944 + 1{,}475 + 26{,}46 + 1364$$

Складывая все эти числа, а затем округлив полученный результат до четырёх значащих цифр, получаем $S = \underline{1393}$.

Однако при вычислении на компьютере округление происходит после каждого сложения. Предполагая условно, что сетка ЭВМ четырёхразрядная, проследим за вычислением на компьютере суммы чисел

Вариант 1: ***от наименьшего к наибольшему***

$0{,}2764 + 0{,}3944 = 0{,}6708$; $0{,}6708 + 1{,}475 = 2{,}156$; $2{,}156 + 26{,}46 = 28{,}62$; $28{,}62 + 1364 = 1393$

получили $S_1 = \underline{1393}$, т.е. верный результат.

Изменим порядок вычислений и начнём складывать числа последовательно

Вариант 2: ***от наибольшего к наименьшему***

$1364+26{,}46 = 1390$; $1390+1{,}475=1391$; $1391+0{,}3944 = 1391$; $1391 + 0{,}2764 = 1391$

получаем $S_2 = \underline{1391}$,

Вывод: ***Необходимо придерживаться правила, в соответствии с которым сложение чисел нужно проводить по мере их возрастания***

Нормализованная форма числа.

Запись числа в форме ***с фиксированной точкой***

$$-0{,}00384 \qquad 345{,}0210$$

Запись числа в форме ***с плавающей точкой***

$$0{,}63750 \cdot 10^6 \qquad 637{,}50 \cdot 10^3 \qquad 6{,}3750 \cdot 10^5$$

Запись числа с плавающей точкой, 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 (кроме нуля целых), т.е. первая цифра после запятой всегда отлична от нуля.***

Такая форма записи числа называется ***нормализованной*** .

В данном примере ею является запись $0{,}63750 \cdot 10^6$

Для числа $-0{,}00384$ нормализованная форма $-0{,}384 \cdot 10^{-2}$.

Итак, общая запись числа a в нормализованной форме имеет вид

$$a = a^0 \cdot 10^p \qquad \text{где} \quad 0{,}1 \le |a^0| < 1.$$

Число a^0 называется ***мантиссой числа***, а число p – его ***порядком***.

Например, для числа $620 = 0{,}620 \cdot 10^3$ мантиссой является $0{,}620$, а порядком – число 3.

Заметим, что ***в этой записи все цифры после запятой верные***.

Задачи теории погрешностей

Прямая задача теории погрешностей

По известным погрешностям некоторой системы параметров требуется определить погрешность функции от этих параметров.

В некоторой области G_n - мерного числового пространства рассматривается непрерывно дифференцируемая функция

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

и пусть Δx^*_i - абсолютные погрешности аргументов. Тогда абсолютная погрешность функции может быть вычислена по формуле Лагранжа

$$\Delta y^* = \left| y - y^* \right| \le \sum_{i=1}^{n} \left| \frac{df(x_1^*, x_2^*, \dots x_n^*,}{dx_i} \right| \Delta x_i^*$$

При зависимости функции от одного параметра $y = f(x)$

$$\Delta y^* \le \left| f'(x^*) \right| \Delta x_i^*$$

Задачи теории погрешностей

Обратная задача теории погрешностей

Определение допустимой погрешности аргументов по допустимой погрешности функции.

Для функции одной переменной $y = f(x)$ абсолютную погрешность можно вычислить приближённо по формуле

$$\boxed{\Delta a = \frac{1}{|f'(a)|} \cdot \Delta y, \qquad f'(a) \neq 0}$$

Для функций нескольких переменных $y = f(x_1, \ldots, x_n)$ задача решается при следующих ограничениях.

- Если значение **одного из аргументов значительно труднее измерить или вычислить** с той же точностью, что и значение остальных аргументов, то погрешность именно этого аргумента и согласовывают с требуемой погрешностью функции.

$$\boxed{\left|\frac{df}{dx_i}\right| \cdot \Delta a_i \qquad i = 1, 2, \ldots, n}$$

- Если значения всех аргументов можно одинаково легко определить с любой точностью, то применяют ***принцип равных влияний***, т.е. учитывают, что все слагаемые

$$\boxed{\Delta a_i = \frac{\Delta y_i}{n} \left|\frac{df}{dx_i}\right|_i \qquad i = 1, 2, \ldots, n}$$

Задачи теории погрешностей

Пример: ***Найти предельные абсолютную и относительную погрешности объема шара*** $V=1/6\pi d^3$, ***если*** $d = 3{,}7\pm0{,}05$ см; $\pi \approx 3{,}14$.

Решение: Рассмотрим d и π как переменные величины.

Вычислим частные производные

$$\frac{\partial V}{\partial d} = \frac{1}{2}\pi d^2, \qquad \frac{\partial V}{\partial \pi} = \frac{1}{6} d^3$$

При заданных значениях d и π получаем, что

$$\frac{\partial V}{\partial d} = \frac{1}{2} 3{,}14\cdot(3{,}7)^2 = 21{,}49; \qquad \frac{\partial V}{\partial \pi} = \frac{1}{6}(3{,}7)^3 = 8{,}44$$

Согласно правилу нахождения предельной абсолютной погрешности, имеем:

$$\Delta V = \left|\frac{\partial V}{\partial d}\right|\cdot|\Delta d| + \left|\frac{\partial V}{\partial \pi}\right|\cdot|\Delta\pi| = 21{,}49\cdot 0{,}05 + 8{,}44\cdot 0{,}0016 = 1{,}101508 \approx 1{,}1\tilde{m}^{3}$$

Поэтому $V \approx 26{,}51\pm1{,}1$ см^3.

Относительная погрешность: $\delta V = 1{,}1/26{,}51 \approx 0{,}04 = 4\%$

Задачи теории погрешностей

Рассмотрим важную задачу теории погрешностей и алгоритм её решения.

Пусть даны приближённые значения a^* и b^* величин с точными значениями a и b соответственно. Определим, какое из равенств $a = a^*$ или $b = b^*$ точнее.

Для решения этой задачи существует следующий алгоритм:

1. Находим приближённые значения a^{**} и b^{**} чисел a и b соответственно, с большим числом знаков после запятой, чем у a^* и b^*

2. находим погрешности вычислений (a^*- a^{**}) и (b^*- b^{**}) как разности между двумя приближёнными значениями чисел a и b соответственно.

3. Определяем предельные абсолютные погрешности вычислений Δa и Δb с избытком (округляем полученные значения погрешности вычислений (a^*- a^{**}) и (b^*- b^{**})).

4. Находим предельные относительные погрешности вычислений по формулам ,

$$\delta(a) = \left(\Delta a /_{|a^*|}\right) 100\%; \qquad \delta(b) = \left(\Delta b /_{|b^*|}\right) 100\%;$$

5. Сравнить предельные относительные погрешности вычисления двух чисел, сделать вывод: если равенство точнее равенства если равенство точнее равенства

Задачи теории погрешностей

Пример. Определить, какое равенство точнее: $\sqrt{44} = 6,63$ или $19/41 = 0,463$

Решение. Здесь $a = \sqrt{44}$; $a^* = 6,63$ и $b = 19/41$; $b^* = 0,463$

1. Найдём приближённые значения чисел a и b с большим числом десятичных знаков:

$a^{**} = 6,633$ $b^{**} = 0,4634$

2. Найдём погрешности вычислений $(a^* - a^{**})$ и $(b^* - b^{**})$:

$(a^* - a^{**}) = 6,63 – 6,633 = -0,003$ $(b^* - b^{**}) = 0,463 – 0,4634 = -0,0004$

3. Определим предельные абсолютные погрешности и с **избытком**:

$\Delta a = 0,004$ $\Delta b = 0,0005$

4. Найдём предельные относительные погрешности вычислений:

$$\delta(a) = (\Delta a / |a^*|)100\% = {}^{0,004}/_{6,63}\ 100\% = 0,06$$

$$\delta(b) = (\Delta b / |b^*|)100\% = {}^{0,0005}/_{0,463}\ 100\% = 0,1$$

5. Поскольку $\delta(a) < \delta(b)$ равенство $\sqrt{44} = 6,63$ точнее равенства $19/41 = 0,463$