1 ВВЕДЕНИЕ В ЧИСЛЕННЫЕ МЕТОДЫ

1.1 Математическое моделирование и вычислительный эксперимент

1.1.1 Вычислительный эксперимент

Вычислительный эксперимент – это технология, основанная на построении и анализе с помощью ЭВМ математических моделей изучаемого объекта. На рисунке 1 представлена схемавычислительного эксперимента.

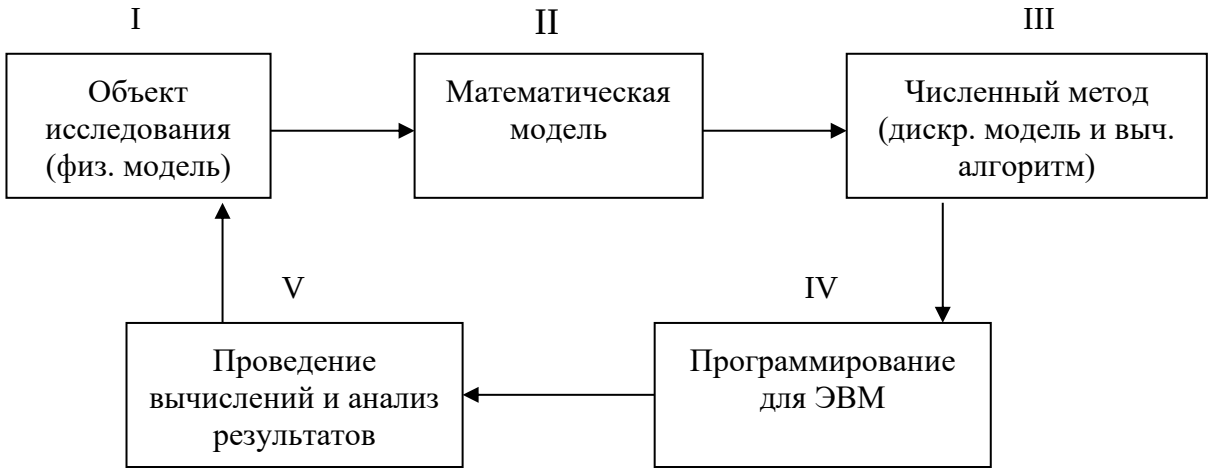


Рисунок 1 – Схема вычислительного эксперимента

Численный метод – это такая интерпретация математической модели («дискретная модель»), которая доступна для реализации на ЭВМ.

Например, если математическая модель представляет собой дифференциальное уравнение, то численным методом может быть аппроксимирующее его разностное уравнение совместно с алгоритмом, позволяющим отыскать решение этого разностного уравнения. Результатом реализации численного метода на ЭВМ является число или таблица чисел.

Необходимо подчеркнуть, что процесс исследования исходного объекта методом математического моделирования и вычислительного эксперимента неизбежно носит приближенный характер, потому что на каждом этапе вносятся те или иные погрешности.

Различают погрешности модели, метода и вычислительные погрешности.

Погрешности модели возникают в связи с упрощением исходного явления, недостаточно точным заданием коэффициентов уравнения и других входных данных.

Погрешности метода – это, как правило**,** погрешности дискретизации.

Погрешность дискретизации возникает при замене непрерывной модели на дискретную, например, при замене дифференциального уравнения системой алгебраических уравнений.

Погрешность округления возникает при задании в ЭВМ входных данных: коэффициентов и правых частей. В процессе работы алгоритма погрешности округления обычно накапливаются, и в результате решение, полученное на ЭВМ, будет отличаться от точного решения дискретизированной задачи. Результирующая погрешность называется погрешностью округления или вычислительной погрешностью. Величина этой погрешности определяется двумя факторами: точностью представления вещественных чисел в ЭВМ и чувствительностью данного алгоритма к погрешностям округления.

Алгоритм называется устойчивым, если в процессе его работы вычислительные погрешности возрастают незначительно и неустойчивым в противном случае.

В общем случае нужно стремиться, чтобы все указанные погрешности имели один и тот же порядок. Например, нецелесообразно пользоваться разностными схемами, имеющими точность 10^{-6}, если коэффициенты исходных уравнении задаются с точностью 10^{-2}.

1.1.2 Требования к вычислительным методам

I группа требований связана с адекватностью дискретной модели исходной математической задаче. К ним относятся: сходимость численного метода, выполнение дискретных аналогов законов сохранения, качественно правильное поведение дискретной задачи.

Известно, что дифференциальные уравнения математической физики являются следствиями интегральных законов сохранения. Поэтому естественно требовать, чтобы для разностной схемы выполнялись аналоги таких законов сохранения.

Сходимость численного метода тесно связана с его корректностью. То есть, если исходная математическая модель поставлена корректно, т.е. ее решение существует, единственно и непрерывно зависит от входных данных, тогда свойство корректности должно сохраняется и для дискретной модели.

II группа требований, предъявляемых к численным методам, связана с возможностью реализации данной дискретной модели на данной ЭВМ. То есть реальные вычислительные алгоритмы должны быть экономичными как по числу операций, так и по требуемому объему памяти.

1.2 Погрешности округления

1.2.1 Представление вещественных чисел в ЭВМ

При ручном счете пользуются десятичной системой счисления, например:

$103{,}67 = 1\cdot10^{2} + 0\cdot10^{1} + 3\cdot10^{0} + 6\cdot10^{-1} + 7\cdot10^{-2}$,

где 10 основание системы счисления. ЭВМ работают в двоичной системе счисления, например:

$0{,}0101=0\cdot2^{0}+0\cdot2^{-1}+1\cdot2^{-2}+0\cdot2^{-3}+1\cdot2^{-4}$

Вообще, в позиционной системе счисления с основанием *r* запись:

$a = \pm a_n a_{n-1}\cdots a_0, a_{-1}a_{-2}\cdots$ означает, что

$$a = \pm (a_n\cdot r^n + a_{n-1}r^{n-1} +\cdots+ a_0\cdot r^0 + a_{-1}\cdot r^{-1} + a_{-2}\cdot r^{-2}+\cdots) \qquad (1)$$

где r – целое число, большее единицы,

$a_i=\{0,1,\cdots r-1\}$

Числа a_i называются разрядами.

Вид (1) называют представлением числа с фиксированной запятой. В ЭВМ чаще всего используется представление чисел в форме с плавающей запятой, в которой любое число представляется в виде:

$$a=Mr^{p}, \qquad (2)$$

где *r* – основание системы счисления, *p* – целое число (положительное, отрицательное, или 0)$r^{-1} \le |M| < 1;$

M называется мантиссой числа *a;*

*P*называется порядком числа *a*.

В виде (2) можно единственным образом представить любое вещественное число кроме нуля.

Например:

$103{,}67 = 0{,}10367\cdot10^{3}$, то есть M=0,10367; p=3

Двоичное число: $0{,}0101=0{,}101\cdot2^{-1}$; M=0,101; p=-1

Так как в ЭВМ для записи каждого числа отводится фиксированное число разрядов (разрядная сетка), следовательно, абсолютная величина чисел будет ограничена. Например, если ЭВМ имеет 48 разрядов, то для записи числа память распределяется следующим образом:

Знак порядка	Порядок (6 разрядов) 111111=63	Знак мантиссы	Мантисса

48 47 p 42 41 40 1 Следовательно,$2^{-63}<|a|\le 2^{63}$ или $10^{-19}<|a|<10^{19}$

1.2.2 Округление чисел в ЭВМ

Минимальное положительное число M_0 которое может быть представлено в ЭВМ с плавающей запятой, называется машинным нулем. Например, для 48 разрядных ЭВМ $M_0 \approx 10^{-19}$. Число $M_\infty = M_0^{-1}$ называется машинной бесконечностью. Все вещественные числа в данной ЭВМ лежат в пределах от M_0до M_∞.

Если в процессе счета какой-либо задачи появится вещественное число, меньшее по модулю чем M_0, то ему присваивается нулевое значение. При появлении в процессе счета вещественного числа большего по модулю, чем M_∞, происходит так называемое переполнение разрядной сетки и ЭВМ прекращает счет задачи. Отметим, что нуль и целые числа представляются в ЭВМ особым образом, так что они могут выходить за пределы диапазона $M_0 \div M_\infty$

Из-за конечности разрядной сетки в ЭВМ можно представить точно не все числа из диапазона $M_0 \div M_\infty$, а лишь конечное множество чисел. Число *a*, не представимое в ЭВМ точно, подвергается округлению $\tilde{a}$. Точность представления характеризуется относительной погрешностью: $\delta = \frac{|a - \tilde{a}|}{|a|}$, величина, которая зависит от способа округления, где числитель этой дроби $\Delta = |a - \tilde{a}|$ называется абсолютной погрешностью.

Простейшим, но не самым точным способом округления является отбрасывание всех разрядов мантиссы числа *a*, которые выходят за пределы разрядной сетки. При таком способе округления граница относительной погрешности имеет вид: $\frac{|a - \tilde{a}|}{|a|} \le 2^{-t+1}$, где *t*– число двоичных разрядов, отводимых для записи мантиссы.

При более точных способах округления можно добиться, чтобы выполнялась оценка: $\frac{|a - \tilde{a}|}{|a|} \le 2^{-t}$.

То есть, относительная точность в ЭВМ с плавающей запятой определяется числом разрядов *t*, отводимых для записи мантиссы. Можно считать, что точное число *a* и отвечающее ему округленное число $\tilde{a}$ связаны равенством: $\tilde{a} = a(1+\varepsilon)$, где $|\varepsilon| \le 2^{-t}$.

Число 2^{-t} называют иногда машинным эпсилоном.Оно характеризует относительную точность представления чисел в ЭВМ.

1.2.3 Накопление погрешностей округления

В процессе вычислений погрешности округления могут накапливаться, так как выполнение каждой из четырех арифметических операций вносит некоторую погрешность. Пусть x – некоторое число, $fl(x)$ – соответствующее ему округленное число. Считается, что выполнение каждой арифметической операции вносит относительную погрешность не большую чем 2^{-t}, то есть

$fl(a * в) = a * в(1+\varepsilon)$ где $* \in \{+,-,\times,\div\}$ и $|\varepsilon| \le 2^{-t}$

Для оценки влияния погрешностей округления на результат того или иного вычисления алгоритма очень часто используют предположение о том, что результат вычислений, измененный погрешностями округления, совпадает с результатом точного выполнения этого же алгоритма, но с иными входными данными.

1.3 Правила приближенных вычислений и оценка погрешностей при вычислениях

1.3.1 Абсолютная погрешность

Под оценкой погрешности приближенного числа aпонимают установление неравенства вида:

$$|a - a_0| \le \Delta_a \tag{3}$$

Число Δ_a называется абсолютной погрешностью приближенного числа a (иногда ее называют «предельная абсолютная погрешность»). Это число определяется неоднозначно: его можно увеличить. Обычно стараются указать возможно меньшее число Δ_a, удовлетворяющее неравенству (3).

Пример 1. Даны длины сторон прямоугольника, измеренные с точностью 0,01. Оценить абсолютную погрешность в определении площади.

$a = 5{,}43; \quad b = 3{,}82; \quad \Delta_a = \Delta_b = 0{,}01$

Решение:

$S = a \cdot b = 20{,}7426$

$S_{\max} = (a + 0{,}01)(b + 0{,}01) = 20{,}8352$

$S_{\min} = (a - 0{,}01)(b - 0{,}01) = 20{,}6502$

$|S - S_{\min}| = 0{,}0924; \quad |S - S_{\max}| = 0{,}0926;$

$\Delta_S = 0{,}0926 \approx 0{,}1$

1.3.2 Относительная погрешность

Относительной погрешностью δ_a приближенного числа a называется отношение: $\delta_a = \dfrac{\Delta_a}{|a|} \qquad a \neq 0$

Относительную погрешность обычно выражают в процентах, и ее принято записывать не более чем с двумя-тремя значащими цифрами (знаками).

Пример 2. Даны длины сторон прямоугольника, измеренные с точностью 0,01. Оценить относительную погрешность в определении площади.

$a = 5{,}43; \quad b = 3{,}82; \quad \Delta_a = \Delta_b = 0{,}01$

Решение: из примера 1 следует, что

$S = a \cdot b = 20{,}7426$

$\Delta_S = 0{,}0926$

следовательно:

$$\delta_S = \frac{0{,}0926}{20{,}7426} = 0{,}0045 = 0{,}45\%$$

Значащие и верные знаки числа.

Значащие цифры числа – это цифры, участвующие в его записи. При подсчете значащих цифр числа не учитывают нулей, стоящих слева, например, в числе 0,010030 пять значащих цифр.

Количество верных знаков числа отсчитывается от первой значащей цифры числа до первой значащей цифры его абсолютной погрешности. Например, если $S = 20{,}7426,\ \ \Delta_S = 0{,}0926$, то S имеет три верных знака (2, 0, 7), остальные знаки сомнительные.

Ориентировочно можно считать, что наличие только одного верного знака соответствует относительной погрешности порядка 10%, двух верных знаков – 1%, трех верных знаков – 0,1% и т.д.

В математических таблицах все числа округлены до верных знаков, причем абсолютная погрешность не превосходит половины единицы последнего оставленного разряда. Например, если в таблице указано $e = 2{,}718$, то абсолютная погрешность не превосходит $0{,}5 \cdot 10^{-3}$.

В окончательных результатах вычислений обычно оставляют кроме верных один сомнительный знак. В промежуточных 2-3 сомнительных знака.