2. Численные методы решения краевых задач

В векторном виде систему ОДУ можно представить в виде

$$\frac{d\mathbf{y}}{dx} = \mathbf{f}(x, \mathbf{y}), a < x < b, \tag{2}$$

Здесь $\mathbf{y}$ и $\mathbf{f}$ - d-мерные вектор-функции. Если (2) дополнить условиями сразу на двух концах в виде

$$\mathbf{g}(\mathbf{y}(a), \mathbf{y}(b)) = \mathbf{0},$$

то такая задача будет называться краевой задачей для ОДУ, а условия - граничными или краевыми.

В качестве примера краевой задачи для ОДУ рассмотрим стационарное уравнение теплопроводности для однородной одномерной среды:

$$u''(x) = f(x), a < x < b, \tag{3}$$

Такое уравнение описывает распределение температуры u в стержне из однородного материала. В качестве граничного условия зададим температуру на обоих концах стержня:

$$u(a) = \alpha, u(b) = \beta \tag{4}$$

Граничное условие (4) называется граничным условием Дирихле. Если вместо самой функции на краях задана ее производная, то такое граничное условие будет называться граничным условием Неймана. Если и то, и другое вместе – граничное условие третьего рода.

Рассмотрим метод численного решения такой задачи. Для простоты будем считать, что $a = 0, b = 1$. Разобъем отрезок $[0, 1]$ с шагом $h = 1/m$ узлами $x_j = jh, j = \overline{0, m}$. На этой сетке в узлах определим сеточную функцию $\{u_i\}_{i=0}^{m}$, u_i - аппроксимация точного решения $u(x_i)$ в i-ом узле сетки. Очевидно, что $u_0 = \alpha, u_m = \beta$ из граничных условий. Аппроксимируем $u^{''}(x)$ в уравнении (3) в каждом внутреннем узле x_j следующим образом:

$$u^{''}(x_j) \approx \frac{1}{h^2}(u_{j-1} - 2u_j + u_{j+1}),$$

тогда для всех внутренних сеточных узлов можно составить следующую систему алгебраических уравнений

$$\frac{1}{h^2}(u_{j-1} - 2u_j + u_{j+1}) = f(x_j), j = \overline{1, m-1} \tag{5}$$

Таким образом, одному дифференциальному уравнению сопоставляется система алгебраических уравнений. Вместе с краевыми условиями они составляют разностную задачу для ОДУ (3) с краевым условием (4). Эту разностную задачу можно представить в матричном виде:

$$AU = F \tag{6}$$

Здесь вектор неизвестных $U = \{u_0, u_1, ..., u_m\}^T$ и

$$A = \frac{1}{h^2}\begin{bmatrix} h^2 & 0 & & & & & \\ 1 & -2 & 1 & & & & \\ & 1 & -2 & 1 & & & \\ & & 1 & -2 & 1 & & \\ & & & \ddots & \ddots & \ddots & \\ & & & & 1 & -2 & 1 \\ & & & & & 0 & h^2 \end{bmatrix}, \ F = \begin{bmatrix} \alpha \\ f(x_1) \\ f(x_2) \\ f(x_3) \\ \dots \\ f(x_{m-1}) \\ \beta \end{bmatrix}.$$

Заметим, что система (6) состоит из $m+1$ уравнения для нахождения $m+1$ неизвестного. Решая полученное СЛАУ, например, методом прогонки или методом простой итерации, находим приближенное решение исходной краевой задачи.

2.1. Локальная и глобальная ошибка численного метода

Построив численные метод и получив решение разностной задачи, остается проверить, насколько оно отличается от решения исходной дифференциальной задачи. Иным словами, оценить ошибку численного метода. Посмотрим, как отличается разностная аппроксимация дифференциального уравнения от самого дифференциального уравнения в произвольном узле j. Для этого подставим в j-ое разностное уравнение системы (5) вместо приближенных значений сеточной функции u_j точные значения решения $u(x_j)$. В общем случае $u(x_j)$ не будут точно удовлетворять разностному уравнению. Это несоответствие называется локальной ошибкой (или невязкой). Обозначим ее τ_j:

$$\tau_j = \frac{1}{h^2}(u(x_{j-1}) - 2u(x_j) + u(x_{j+1})) - f(x_j). \qquad (7)$$

Понятно, что локальную ошибку можно записать так для любого сеточного узла (для любого уравнения). На практике мы не знаем точного решения $u(x_j)$, но используя разложение в ряд Тейлора, можно переписать (7) в виде:

$$\tau_j = [u^{''}(x_j) + \frac{1}{12}h^2 u^{''''}(x_j) + O(h^4)] - f(x_j). \qquad (8)$$

Используя исходное уравнение (3), получим

$$\tau_j = \frac{1}{12}h^2 u^{''''}(x_j) + O(h^4)$$

Из этого уравнения видно, что $\tau_j = O(h^2)$. Обозначим τ - вектор с компонентами τ_j, $\hat{U}$ - вектор точного решения. Тогда

$$\tau = A\hat{U} - F$$

или

$$A\hat{U} = F + \tau \qquad (9)$$

Назовем глобальной ошибкой численного метода вектор $E = U - \hat{U}$. Тогда, используя (6) и (9), можно получить

$$AE = -\tau \qquad (10)$$

Оценим глобальную ошибку E. Для этого перепишем (10)

$$A^h E^h = -\tau^h \qquad (11)$$

Надстрочный индекс h указывает здесь на то, что используется сетка с шагом h. Домножим эту систему на $(A^h)^{-1}$ и возьмем норму:

$$||E^h|| = ||(A^h)^{-1}\tau^h|| \leq ||(A^h)^{-1}||||\tau^h|| \qquad (12)$$

Мы знаем, что норма невязки $||\tau^h|| = O(h^2)$. Чтобы тоже самое можно было сказать о норме глобальной ошибки $||E^h||$, необходимо, чтобы $||(A^h)^{-1}||$ была ограничена некоторой константой C, не зависящей от h при $h \to 0$. Если это так, то справедливо

$$||E^h|| \leq C||\tau^h|| \qquad (13)$$

То есть $||E^h||$ стремится к нулю с тем же порядком, что и $||\tau^h||$.

2.2. Аппроксимация, устойчивость, сходимость

Подводя итог всему вышесказанному, зафиксируем:

1. Численный метод *аппроксимирует* исходную дифференциальную задачу, если $||\tau^h|| \to 0$ при $h \to 0$. Численный метод аппроксимирует исходную задачу с порядком p, если $||\tau^h|| = O(h^p)$

2. Численный метод для линейной краевой задачи *устойчив*, если $||(A^h)^{-1}||$ ограничена некоторой константой , не зависящей от h при $h \to 0$.

3. Численный метод *сходится* к решению исходной дифференциальной задачи, если $||E^h|| \to 0$ при $h \to 0$. При этом говорят, что численный метод сходится с порядком p, если $||E^h|| \leq ||h^p||$.

4. Согласно основной теореме вычислительной математики, численный метод *сходится* к решению исходной дифференциальной задачи, если он ее аппроксимирует и устойчив. Если численный метод аппроксимирует исходную задачу с порядком p, то и сходимость так же будет порядка p.

Замечание: до постороения численного метода прежде стоит задуматься, корректна ли постановка данной дифференциальной задачи. Чтобы численный метод сходился к чему надо, необходимо, чтобы решение задачи существовало и было единственным. В частности, если для (3) использовать граничные условия Неймана, решение не будет единственным.

2.3. Метод построения общего решения

В начале рассмотрим систему ОДУ первого порядка:

$$\frac{d\mathbf{u}(x)}{dx} = \mathbf{G}(x)\mathbf{u} + \mathbf{f}(x),\ x \in [a, b] \tag{14}$$

$\mathbf{u}$, $\mathbf{f}$ - векторы размерности n, $\mathbf{G}$ - матрица размера $n \times n$.

Если матрица $\mathbf{G}$ постоянная и имеет базис из собственных векторов $\mathbf{h}_1, \ldots, \mathbf{h}_n$ (с соответствующими собственными числами $\lambda_1, \ldots, \lambda_n$), то общее решение линейной однородной системы

$\frac{d\mathbf{u}(x)}{dx} = \mathbf{G}\mathbf{u}(x)$ имеет вид:

$$\mathbf{u}(x) = C_1 \exp(\lambda_1 x)\mathbf{h}_1 + \ldots + C_n \exp(\lambda_n x)\mathbf{h}_n$$

Для выделения единственного решения необходимо задать n краевых условий, ограничимся пока рассмотрением условий такого вида:

$$\mathbf{A}\mathbf{u}(a) + \mathbf{B}\mathbf{u}(b) = \mathbf{c} \tag{15}$$

где $\mathbf{A}$,$\mathbf{B}$ – матрицы $n \times n$, $\mathbf{c}$ – вектор длины n. При это предполагается, что краевые условия линейно независимы:

$$\mathrm{rank}[\mathbf{A}, \mathbf{B}] = n \tag{16}$$

Из теории ОДУ известно, что общее решение (14) представляется в виде:

$$\mathbf{u}(x) = \mathbf{u}^{(0)}(x) + \sum_{i=1}^{n} \alpha_i \mathbf{u}^{(i)}(x) \tag{17}$$

где $\mathbf{u}^{(0)}(x)$ – произвольное решение неоднородной системы

$$(\mathbf{u}^{(0)})_x = \mathbf{G}(x)\mathbf{u}^{(0)} + \mathbf{f}(x)$$

а $\mathbf{u}^{(i)}(x)$ – n линейно независимых решений однородной системы

$$(\mathbf{u}^{(i)})_x = \mathbf{G}(x)\mathbf{u}^{(i)}$$

Подбирая n свободных коэффициентов α_i мы можем удовлетворить n краевых условий (15). Подставив общее решение в краевые условия, получим систему *линейных алгебраических уравнений* относительно α_i:

$$\mathbf{A}\left[\mathbf{u}^{0}(a) + \sum_{i=1}^{n} \alpha_i \mathbf{u}^{(i)}(a)\right] + \mathbf{B}\left[\mathbf{u}^{(0)}(b) + \sum_{i=1}^{n} \alpha_i \mathbf{u}^{(i)}(b)\right] = \mathbf{c} \tag{18}$$

Если данная система имеет единственное решение, краевая задача имеет единственное решение.

Используем теперь эти теоретические результаты для численного решения краевой задачи. В предыдущих разделах были изложены методы решения задачи Коши для системы (14). Применим их:

- решение $\mathbf{u}^{(0)}(x)$ получим, решая задачу Коши для системы (14) с начальными данными $\mathbf{u}^{(0)}(a) = 0$ каким-либо известным методом (например, явным методом Эйлера, или методом Рунге-Кутты).

- для вычисления линейно-независимых решений $\mathbf{u}^{(i)}(x)$ тем же численным методом решим задачи Коши для системы $(\mathbf{u}^{(i)})_x = \mathbf{G}(x)\mathbf{u}^{(i)}$ с начальными данными $\mathbf{u}^{(i)}(a) = \mathbf{e}^{(i)} = [0, \ldots, 0, 1_i, 0, \ldots,])^T$ (на i-ом месте стоит единица). Такие начальные данные обеспечивают линейную независимость решений $\mathbf{u}^{(i)}(x)$

- Используя найденные значения в точке b, решим линейную систему (18) и найдем коэффициенты α_i.

Получается, что для решения краевой задачи (14) + (15) необходимо численно решить $n+1$ задачу Коши.

Замечание: данный метод действительно позволяет построить общее решение: решив один раз задачи Коши, мы можем находить решения разных краевых задач (для разных краевых условий), только лишь решая одну систему линейных алгебраических уравнений относительно коэффициентов α_i

Задача 5 (общее решение краевой задачи). *Численно найти решение краевой задачи* $y'' + 10xy' - 2y = 1$, $0 \le x \le 2$, $y(0) = \alpha, y(2) = \beta$ *на сетке* $\{0, 1, 2\}$.

Решение: общее решение представимо в виде $y(x) = c_1 u(x) + c_2 v(x) + w(x)$, где u, v и w находятся из решения следующих задач Коши:

$$\begin{cases} u'' + 10xu' - 2u = 0, 0 \le x \le 2, \\ u(0) = 0, u'(0) = 1. \end{cases}$$

$$\begin{cases} v'' + 10xv' - 2v = 0, 0 \le x \le 2, \\ v(0) = 1, v'(0) = 0. \end{cases}$$

$$\begin{cases} w'' + 10xw' - 2w = 1, 0 \le x \le 2, \\ w(0) = 0, w'(0) = 0. \end{cases}$$

Чтобы решить эти задачи, например, явным методом Эйлера, сведем каждую из задач к системе первого порядка. Рассмотрим на примере задачи на u. Пусть $u = u^1$, $u' = u^2$. Тогда

$$\begin{cases} (u^1)' = u^2, u^1 = 0, 0 \le x \le 2, \\ (u^2)' = -10xu^2 + 2u^1, u^2 = 1, 0 \le x \le 2. \end{cases}$$

Решая численно задачу на сетке из трех узлов и возвращаясь к переменной u, получим $u = (0, 1, 2)^T$. Аналогично и для других переменных $v = (1, 1, 3)^T$, $w = (0, 0, 1)^T$. Решение задачи $y = c_1(0, 1, 2)^T + c_2(1, 1, 3)^T + (0, 0, 1)^T$ ■

2.4. Конечно-разностный метод решения линейного уравнения второго порядка

Рассмотрим краевую задачу для линейного уравнения второго порядка:

$$-u'' + p(x)u' + q(x)u = r(x),\ x \in [a, b] \tag{19}$$

для начала, с простыми краевыми условиями:

$$u(a) = \alpha, u(b) = \beta \tag{20}$$

Если p, q, r непрерывны и $q > 0$ на $[a, b]$, то задача (19) + (20) имеет единственное решение.

В уравнение входят вторая и первая производная функции u, аппроксимируем их конечными разностями. Введём сетку $\{x\}_{j=0}^{m+1}$, $h = \frac{(b-a)}{m+1}$. Обозначим через u_j значения в узлах некоторой сеточной функции, которая, как мы ожидаем, будет приближать значения точного решения $u(x_j)$ в узлах. Во всех внутренних узлах $(j = \overline{1, m})$ заменим первую и вторую производные конечно-разностными аппроксимациями второго порядка, а для граничных узлов добавим очевидные равенства:

$$\begin{aligned} &u_0 = \alpha \\ &-\frac{u_{j+1} - 2u_j + u_{j-1}}{h^2} + p_j \frac{u_{j+1} - u_{j-1}}{2h} + q_j u_j = r_j,\ j = 1, \ldots, m \\ &u_{m+1} = \beta \end{aligned} \tag{21}$$

Если обозначить $U = u_0, u_1, \ldots, u_{m+1}{}^T$, полученную систему можно переписать в виде:

$$AU = F \tag{22}$$

где A – матрица $m + 2 \times m + 2$, имеющая трёхдиагональную структуру.

$$A = \frac{1}{h^2}\begin{bmatrix} h^2 & 0 & & & & \\ -1-\frac{p_1}{2h} & 2+h^2q_1 & -1-\frac{p_1}{2h} & & & \\ & -1-\frac{p_2}{2h} & 2+h^2q_2 & -1-\frac{p_2}{2h} & & \\ & & \ddots & \ddots & \ddots & \\ & & & -1-\frac{p_m}{2h} & 2+h^2q_m & -1-\frac{p_m}{2h} \\ & & & & 0 & h^2 \end{bmatrix}, \tag{23}$$

$F = [\alpha, r_1, r_2, \ldots, r_m, \beta]^T$.

Замечание: на одном из прошлых семинаров рассматривалась краевая задача для более простого уравнения $u''(x) = f(x)$. Обратите внимание, что в том случае матрица имела более простой вид, что сильно упрощало анализ её свойств.

После того, как разностные уравнения записаны в виде системы (22), возникает два вопроса:

1. Имеет ли система решение? Является ли матрица невырожденной? То есть можно ли из этих уравнений однозначно определить вектор приближённого решения?

2. Приближает ли приближённое решение точное?

Ответ на первый вопрос: если $h < 2/L$, где $L = max_{x\in[a,b]}|p(x)|$, линейная система (22) имеет единственное решение. Второй вопрос снова приводит нас к проблеме исследования сходимости.

2.5. Граничные условия Неймана

Помимо простых краевых условий на значения искомой функции на краях интервала (граничные условия Дирихле), в некоторых задачах на одной из границ краевое условия ставится для производной:

$$u'(a) = \sigma, u(b) = \beta \tag{24}$$

В задаче о распределении температуры в одномерном стержне это соответствует заданному тепловому потоку на одном из концов. Например, если один из концов стержня теплоизолирован, то тепловой поток на этом конце должен быть равен нулю.

Теперь необходимо аппроксимировать и производную в граничном условии.

Первый подход Можно взять вместо первой производной одностороннюю аппроксимацию:

$$\frac{u_1 - u_0}{h} = \sigma \tag{25}$$

Однако, в этом случае схема будет аппроксимировать задачу с только первым порядком.

Второй подход Чтобы получить метод второго порядка, нужно использовать центральную разностную формулу для первой производной. Для этого удобно добавить «искусственный» узел и значение в нём u_{-1}. Вместо одного разностного уравнения (25) получатся два уравнения:

$$-\frac{u_1 - 2u_0 + u_{-1}}{h^2} + p_0 \frac{u_1 - u_{-1}}{2h} + q_0 u_0 = r_0 \tag{26}$$

$$\frac{u_1 - u_{-1}}{2h} = \sigma \tag{27}$$

Введение узла вне области определения дифференциальной задачи может показаться странным, однако, можно просто исключить u_{-1} и получить одно уравнение:

$$\frac{u_1 - u_0}{h} = \sigma + \frac{h}{2}(p_0 \sigma + q_0 u_0 - r_0) \tag{28}$$

Можно истолковать эту формулу как одностороннюю формулу с некоторой поправкой (по сути, вычитается главный член ошибки аппроксимации, чтобы получить второй порядок аппроксимации).

Третий подход Вместо центральной формулы второго порядка, можно использовать одностороннюю формулу второго порядка:

$$\frac{1}{h}\left(\frac{3}{2}u_0 - 2u_1 + \frac{1}{2}u_2\right) = \sigma \tag{29}$$

2.6. Решение линейных систем

Как мы убедились, при аппроксимации производных дискретными соотношениями, дифференциальный линейный оператор аппроксимируется оператором, действующим на элемент конечно-мерного пространства (вектор), т.е. дифференциальный оператор заменяется некоторой матрицей. Важно отметить, что как данном случае, так и в дальнейшем, матрицы линейных систем, возникающие в численных методах, имеют определённую структуру, что позволяет эффективно решать линейные системы очень большой размерности.

Метод прогонки

Для решения линейных системы с трёхдиагональной матрицей можно использовать метод прогонки. Он позволяет решить систему размерности N за $O(N)$ операций. Перепишем уравнения в таком виде:

$$\begin{aligned} &b_1u_1 + c_1u_2 = d_1 \\ &a_nu_{n-1} + b_nu_n + c_nu_{n+1} = d_n,\ n = 1, \ldots, N-1 \\ &a_Nu_{N-1} + b_Nu_N = d_N \end{aligned} \tag{30}$$

Будем искать решение в виде прогоночного соотношения

$$u_{n-1} = A_nu_n + F_n \tag{31}$$

Первое уравнение можно переписать в виде:

$$u_1 = -\frac{c_1}{b_1}u_2 + \frac{d_1}{b_1} \tag{32}$$

т.е. $A_1 = -\frac{c_1}{b_1}, F_1 = \frac{d_1}{b_1}$ Подставив равенство $u_{n-1} = A_nu_n + F_n$ в уравнение $a_nu_{n-1} + b_nu_n + c_nu_{n+1} = d_n$, получим:

$$u_n = -\frac{c_n}{a_nA_n + b_n}u_{n+1} + \frac{d_n - a_nF_n}{a_nA_n + b_n}$$

Сравнивая полученное выражение с $u_{n-1} = A_n u_n + F_n$ приходим к выводу, что

$$A_{n+1} = -\frac{c_n}{a_n A_n + b_n}, F_{n+1} = \frac{d_n - a_n F_n}{a_n A_n + b_n}$$

Эти формулы определяют *прямой ход прогонки*: последовательно находим все коэффициенты A_n, F_n и запоминаем их.
Далее, из последнего прогоночного соотношения $u_{N-1} = A_N u_N + F_N$ и последнего уравнения системы $a_N u_{N-1} + b_N u_N = d_N$ находим величины u_N и u_{N-1} (решаем линейную систему 2×2). После этого из прогоночных соотношений 31 последовательно вычисляем $u_n, n = N-2, \ldots, 1$ - это *обратный ход прогонки.*

Следующее условие гарантирует устойчивость метода прогонки:
$|b_n| > |a_n| + |c_n|$. Это условие называется *условием диагонального преобладания.*