

1. Связь метода Гаусса с факторизацией матрицы

Рассмотрим матричное уравнение вида

$$Ax = f, \tag{1}$$

где $|A| \neq 0$, A $(m \times m)$, $x = (x_1, x_2, \dots, x_m)^T$, $f = (f_1, f_2, \dots, f_m)^T$. Матрица A , вообще говоря, может быть матрицей с комплексными элементами.

Рассмотрим факторизацию (разложение в произведение) матрицы A $(m \times m)$

$$A = B \cdot C, \tag{2}$$

где B — нижняя треугольная матрица, а C — верхняя треугольная матрица с единицами на главной диагонали:

$$B = \begin{pmatrix} b_{11} & 0 & \cdots & 0 \\ b_{21} & b_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mm} \end{pmatrix}, \quad C = \begin{pmatrix} 1 & c_{12} & \cdots & c_{1m} \\ 0 & 1 & \cdots & c_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$

$$b_{ij} = a_{ij} - \sum_{l=1}^{j-1} b_{il}c_{lj}, \quad i \geqslant j. \quad c_{ij} = \frac{a_{ij} - \sum_{l=1}^{i-1} b_{il}c_{lj}}{b_{ii}}, \quad i < j. \tag{3}$$

Утверждение. Пусть все угловые миноры матрицы A отличны от нуля. Тогда представление матрицы A в виде (2) существует и единственно.

Задача. Показать, что для вычисления элементов матриц B и C по формулам (3) требуется $\frac{m^3-m}{3}$ умножений и делений. (Число умножений и делений далее будем называть числом операций.)

Замечание. Число действий, необходимое для преобразований матрицы в прямом ходе метода Таким образом факторизация матрицы A в виде (2) требует такое же число действий, что и сведение матрицы A к A' в прямом ходе метода Гаусса.

$$\begin{cases} By = f \\ Cx = y, \quad y = (y_1, \dots, y_m)^T. \end{cases} \tag{4}$$

$$\tag{5}$$

Замечание 1. На вычисление новых правых частей, т.е. вектора f' , в методе Гаусса уходит $\frac{m(m+1)}{2}$ действий.

Замечание 2. Число операций, затрачиваемых на выполнение обратного хода метода Гаусса, равно $\frac{m(m-1)}{2}$, что совпадает с числом действий, требуемых для решения второй системы из (4).

В итоге получим, что для решения систем (4) требуется $\frac{m(m-1)}{2} + \frac{m(m+1)}{2} = m^2$ операций. Тогда все решение системы (1) с использованием факторизации матриц требует $\frac{m^3-m}{3} + m^2 = \frac{m^3+3m^2-m}{3}$ операций

2. Обращение матрицы методом Гаусса-Жордана

Рассмотрим задачу обращения (поиска обратной матрицы) невырожденной матрицы A $(m \times m)$.

$$AX = E, \quad \sum_{l=1}^m a_{il}x_{lj} = \delta_{ij}, \quad i = \overline{1, m}, \quad j = \overline{1, m}. \tag{6}$$

где A $(m \times m)$, $|A| \neq 0$, при этом количество операций равно

$$\frac{m^3-m}{3} + \frac{m(m+1)(m+2)}{6} + \frac{m^2(m-1)}{2} = m^3.$$

Отметим, что он является самым эффективным методом обращения невырожденных матриц произвольного вида.

3. Метод квадратного корня решения СЛАУ

Определение. Квадратная матрица A называется эрмитовой (самосопряженной), если ее элементы связаны соотношением $a_{ij} = \overline{a_{ji}}$. В этом случае будем записывать $A = A^*$.

Факторизуем эрмитову матрицу A в виде

$$A = S^*DS, \tag{7}$$

где матрица S — верхнетреугольная матрица с положительными элементами на главной диагонали, а D — диагональная матрица со значениями ± 1 на главной диагонали:

$$S = \begin{pmatrix} s_{11} & s_{12} & \cdots & s_{1m} \\ 0 & s_{22} & \cdots & s_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & s_{mm} \end{pmatrix}, \quad s_{ii} > 0,$$

$$D = \begin{pmatrix} d_{11} & 0 & \cdots & 0 \\ 0 & d_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_{mm} \end{pmatrix}, \quad d_{ii} = \pm 1.$$

Выразим d_{ii} и s_{ii} :

$$d_{ii} = \operatorname{sgn}(a_{ii} - \sum_{l=1}^{i-1} |s_{li}|^2 d_{ll}), \quad i = \overline{1, m}, \tag{8}$$

$$s_{ij} = \frac{a_{ij} - \sum_{l=1}^{i-1} \overline{s_{li}} d_{ll} s_{lj}}{s_{ii} d_{ii}}, \quad i, j = \overline{1, m}, \quad i < j. \tag{9}$$

Метод квадратного корня позволяет примерно вдвое уменьшить число операций, необходимых для решения системы, по сравнению с методом Гаусса — до $\sim \frac{m^3}{6}$ умножений и делений. Кроме этого необходимо m операций извлечения квадратного корня. Заметим, что метод справедлив только в случае, если матрица системы линейных уравнений эрмитова.

4. Примеры и канонический вид итерационных методов решения СЛАУ

Определение. Итерационный метод называется двухслойным, если для вычисления текущей итерации используются только элементы предыдущей итерации.

Замечание. Двухслойный итерационный метод также называют одношаговым.

Метод Якоби

Метод Якоби является явным итерационным методом и задается правилом

$$x_i^{n+1} = \frac{f_i - \sum_{j=1}^{i-1} a_{ij}x_j^n - \sum_{j=i+1}^m a_{ij}x_j^n}{a_{ii}}, \quad i = \overline{1, m}, \quad n \in \mathbb{Z}_+.$$

Метод Якоби является легко реализуемым, но при этом медленно сходящимся, особенно при больших m .

Метод Зейделя

$$x_i^{n+1} = \frac{f_i - \sum_{j=1}^{i-1} a_{ij}x_j^{n+1} - \sum_{j=i+1}^m a_{ij}x_j^n}{a_{ii}}, \quad i = \overline{1, m}, \quad n \in \mathbb{Z}_+.$$

В правой части уравнения используются координаты $(n+1)$ -й итерации, поэтому метод Зейделя является неявным. Но если разумно организовать вычисления, то можно найти координаты $(n+1)$ -й итерации по явным формулам.

Зейделя прост в реализации, но медленно сходится.

Каноническая запись итерационных методов

$$\text{Метод Якоби:} \quad D(x^{n+1} - x^n) + Ax^n = f, \quad n \in \mathbb{Z}_+, \tag{10}$$

$$\text{Метод Зейделя:} \quad (D + R_1)(x^{n+1} - x^n) + Ax^n = f, \quad n \in \mathbb{Z}_+. \tag{11}$$

Определение. Канонической формой записи двухслойного итерационного метода решения системы называется его запись в виде

$$B_{n+1} \frac{x^{n+1} - x^n}{\tau_{n+1}} + Ax^n = f, \tag{12}$$

где $n \in \mathbb{Z}_+$, начальное приближение x^0 задано, τ_{n+1} — положительное вещественное число, называемое итерационным параметром, B_{n+1} — некоторая обратимая матрица.

Определение. Если в методе (12) параметр τ_{n+1} и матрица B_{n+1} не зависят от номера итерации ($B_{n+1} = B$, $\tau_{n+1} = \tau$), то такой метод называется стационарным, в противном случае — нестационарным.

Определение. Если $B_{n+1} = E$, то метод (12) называется явным, в противном случае — неявным.

При рассмотрении итерационных методов обычно исследуют условия, при которых данный метод сходится, и оценивают скорость сходимости метода.

Метод простой итерации

Метод простой итерации (метод релаксации) определяется итерационной схемой вида

$$\frac{x^{n+1} - x^n}{\tau} + Ax^n = f, \quad \tau > 0, \quad n \in \mathbb{Z}_+, \quad x^0 \text{ — задано.} \tag{13}$$

Метод Ричардсона

Метод Ричардсона определяется итерационной схемой вида

$$\frac{x^{n+1} - x^n}{\tau_{n+1}} + Ax^n = f, \quad \tau_{n+1} > 0, \quad n \in \mathbb{Z}_+, \quad x^0 \text{ — задано.} \tag{14}$$

Замечание. Для итерационных методов (13) и (14) в случае, когда матрица A является симметричной и положительно определенной, известен такой набор итерационных параметров (Чебышевский набор), при котором сходимость этих методов будет наиболее быстрой.

Попеременно-треугольный итерационный метод

Представим матрицу A в виде $A = R_1 + R_2$, где R_1 — нижнетреугольная матрица, R_2 — верхнетреугольная матрица:

$$R_1 = \begin{pmatrix} 0.5a_{11} & 0 & \cdots & 0 \\ a_{21} & 0.5a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & 0.5a_{mm} \end{pmatrix},$$
$$R_2 = \begin{pmatrix} 0.5a_{11} & a_{12} & \cdots & a_{1m} \\ 0 & 0.5a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0.5a_{mm} \end{pmatrix}.$$

Попеременно-треугольный метод имеет вид

$$(E + \omega R_1)(E + \omega R_2) \frac{x^{n+1} - x^n}{\tau} + Ax^n = f, \quad n \in \mathbb{Z}_+, \tag{15}$$

где $\tau > 0$, $\omega > 0$ — итерационные параметры

Определение. Вектор $r^n = f - Ax^n$ называется невязкой на n -й итерации.

5. Теорема Самарского о сходимости итерационных методов

Рассмотрим матричное уравнение вида

$$Ax = f, \tag{16}$$

где $|A| \neq 0$, $A \ (m \times m)$, $x = (x_1, x_2, \dots, x_m)^T$, $f = (f_1, f_2, \dots, f_m)^T$. Рассмотрим также двухслойный стационарный метод решения уравнения (16):

$$B \frac{x^{n+1} - x^n}{\tau} + Ax^n = f, \tag{17}$$

где $n \in \mathbb{Z}_+$, начальное приближение x^0 задано, τ — положительное вещественное число, B — обратимая матрица порядка $(m \times m)$.

Введем евклидову норму: $\|x\| = \sqrt{(x, x)} = \left(\sum_{i=1}^m x_i^2\right)^{\frac{1}{2}}$. Эту норму также часто называют среднеквадратичной нормой.

Определение. Линейный оператор D называется положительным (неотрицательным), если $(Dx, x) > 0 \ \forall x \in H$, $x \neq \theta$ (соответственно $(Dx, x) \geq 0 \ \forall x \in H$). Положительность оператора D обозначается как $D > 0$.

Определение. Скалярным произведением в смысле оператора D называется скалярное произведение, определяемое соотношением $(x, y)_D = (Dx, y)$.

Определение. Энергетической нормой, порождаемой линейным самосопряженным положительно определенным оператором D , называется норма, задаваемая соотношением $\|x\|_D = \sqrt{(x, x)_D} = \sqrt{(Dx, x)}$.

Задача. Пусть $D = D^* > 0$. Доказать, что $\exists \delta > 0 : (Dx, x) \geq \delta(x, x) = \delta\|x\|^2$.

Рассмотрим свойства положительного самосопряженного линейного оператора.

Если $D = D^* > 0$, то определены $D^{-1} = \left(D^{-1}\right)^* > 0$, $D^{\frac{1}{2}} = \left(D^{\frac{1}{2}}\right)^* > 0$, $D^{-\frac{1}{2}} = \left(D^{-\frac{1}{2}}\right)^* > 0$.

Определение. Погрешностью итерационного метода на n -й итерации называется вектор

$$v^n = x^n - x. \tag{18}$$

Определение. Итерационный метод сходится в норме $\|\cdot\|$, если $\|v^n\| \rightarrow 0$ при $n \rightarrow \infty$.

Определение. Матрица $S = E - \tau B^{-1}A$ называется матрицей перехода от n -й итерации к $(n + 1)$ -й.

Теорема 1. Итерационный метод (17) решения системы (16) сходится при любом начальном приближении тогда и только тогда, когда все собственные значения матрицы перехода S по модулю меньше единицы. (Без доказательства)

Заметим, что данная теорема практически неприменима, так как задача нахождения полного спектра матрицы S аналитически решается крайне редко.

Теорема 2 (теорема Самарского). Пусть A — самосопряженный положительно определенный оператор, τ — положительное вещественное число и выполнено операторное неравенство

$$B - \frac{\tau}{2}A > 0. \tag{19}$$

Тогда итерационный метод (17) решения системы (16) сходится в среднеквадратичной норме при любом начальном приближении:

$$\|x^n - x\| = \sqrt{\sum_{j=1}^m \left(x_j^n - x_j\right)^2} \xrightarrow{n \rightarrow \infty} 0, \quad \forall x^0.$$

6. Достаточные условия сходимости методов Якоби, Зейделя, простой итерации.

Теорема 3 (теорема Самарского). Пусть A — самосопряженный положительно определенный оператор, τ — положительное вещественное число и выполнено операторное неравенство

$$B - \frac{\tau}{2}A > 0. \tag{20}$$

Тогда итерационный метод (17) решения системы (16) сходится в среднеквадратичной норме при любом начальном приближении:

$$\|x^n - x\| = \sqrt{\sum_{j=1}^m \left(x_j^n - x_j\right)^2} \xrightarrow{n \rightarrow \infty} 0, \quad \forall x^0.$$

Следствие 1. Пусть $A = A^* > 0$. Тогда метод Якоби сходится в среднеквадратичной норме при любом начальном приближении, если выполнено неравенство:

$$2D > A,$$

где $A = R_1 + D + R_2$, $D = \text{diag}(a_{11}, a_{22}, \dots, a_{mm})$.

Следствие 2. Пусть положительная симметричная матрица ($A = A^* > 0$) является матрицей со строгим диагональным преобладанием:

$$a_{ii} > \sum_{j=1, j \neq i}^m |a_{ij}|, \quad i = \overline{1, m}.$$

Тогда метод Якоби сходится в среднеквадратичной норме при любом начальном приближении x^0 .

Задача. Пусть матрица $A = A^* > 0$. Доказать, что $a_{ii} > 0$, $i = \overline{1, m}$.

Следствие 3. Пусть $A = A^* > 0$. Тогда метод Зейделя сходится в среднеквадратичной норме при любом начальном приближении x^0 .

Следствие 4. Пусть $A = A^* > 0$, $\gamma_2 = \max_{1 \leq k \leq m} \lambda_k > 0$. Если $0 < \tau < \frac{2}{\gamma_2}$, то метод простой итерации сходится в среднеквадратичной норме при любом начальном приближении x^0 .

7. Теорема об оценке скорости сходимости итерационных методов.

Для достижения заданной точности ϵ достаточно провести число итераций, равное $n_0(\epsilon) = \left[\frac{\ln \frac{1}{\epsilon}}{\ln \frac{1}{\rho}}\right]$, где $[x]$ — целая часть числа x .

Определение. Величина $\ln \frac{1}{\rho}$ называется скоростью сходимости итерационного метода.

Пусть $D = D^* > 0$. Введем энергетическую норму, порождаемую оператором D : $\|x\|_D = \sqrt{(Dx, x)}$.

В пространстве H существует ортонормированный базис $\{e_k\}$ из собственных векторов оператора D .

Теорема 4 (об оценке скорости сходимости). Пусть $A = A^* > 0$, $B = B^* > 0$. Пусть также существует число ρ , $0 < \rho < 1$, такое, что выполнено операторное неравенство:

$$\frac{1 - \rho}{\tau}B \leq A \leq \frac{1 + \rho}{\tau}B. \tag{21}$$

Тогда для погрешности итерационного метода $B \frac{x^{n+1} - x^n}{\tau} + Ax^n = f$ решения системы $Ax = f$ справедлива оценка:

$$\|v^{n+1}\|_B \leq \rho \|v^n\|_B, \quad n \in \mathbb{Z}_+. \tag{22}$$

Замечание. Оценка (22) справедлива и в энергетической норме $\|\cdot\|_A$.

Следствие 5. Пусть A, B — самосопряженные положительно определенные операторы, и пусть существуют $\gamma_2 > \gamma_1 > 0$, для которых выполняется условие $\gamma_1 B \leq A \leq \gamma_2 B$. Тогда, если $\tau = \tau_0 = \frac{2}{\gamma_1 + \gamma_2}$, то двухслойный итерационный метод решения системы уравнений сходится, и верна оценка

$$\|x^{n+1} - x\|_B \leq \rho \|x^n - x\|_B, \tag{23}$$

где $\rho = \frac{1 - \xi}{1 + \xi}$, $\xi = \frac{\gamma_1}{\gamma_2}$.

Следствие 6. Для метода простой итерации, пусть A — самосопряженный положительно определенный оператор, а γ_1 и γ_2 — его минимальное и максимальное собственные значения: $\gamma_1 = \min_{1 \leq k \leq m} \lambda_k^A$, $\gamma_2 = \max_{1 \leq k \leq m} \lambda_k^A$. Кроме того, пусть $\tau = \frac{2}{\gamma_1 + \gamma_2}$. Тогда верна оценка $\|x^{n+1} - x\| \leq \rho \|x^n - x\|$, где $\rho = \frac{1 - \xi}{1 + \xi}$, $\xi = \frac{\gamma_1}{\gamma_2}$.

8. Попеременно-треугольный итерационный метод. Реализация метода. Теорема о сходимости.

Запишем каноническую форму попеременно-треугольного итерационного метода (ПТИМ):

$$(E + \omega R_1)(E + \omega R_2) \frac{x^{n+1} - x^n}{\tau} + Ax^n = f, \quad \omega > 0, \quad \tau > 0, \quad n \in \mathbb{Z}_+.$$

Обозначим $B = (E + \omega R_1)(E + \omega R_2)$.

Теорема 5 (о сходимости ПТИМ). Пусть A — самосопряженный положительно определенный оператор и $\omega > \frac{1}{4}$. Тогда ПТИМ сходится в среднеквадратичной норме при любом начальном приближении x^0 .

9. Теорема об оценке скорости сходимости попеременно-треугольного итерационного метода

Теорема 6 (о скорости сходимости ПТИМ). Пусть A — самосопряженный положительно определенный оператор и числа $\delta > 0$, $\Delta > 0$ таковы, что выполняются неравенства

$$A \geq \delta E, \quad R_2^* R_2 \leq \frac{\Delta}{4} A. \tag{24}$$

Положим $\omega = \frac{2}{\sqrt{\delta\Delta}}$, $\tau = \frac{2}{\gamma_1 + \gamma_2}$, $\gamma_1 = \frac{\sqrt{\delta}}{2} \left(\frac{\sqrt{\delta\Delta}}{\sqrt{\delta} + \sqrt{\Delta}} \right)$, $\gamma_2 = \frac{\sqrt{\delta\Delta}}{4}$. Тогда ПТИМ сходится и имеет место оценка $\|x^{n+1} - x\|_B \leq \rho \|x^n - x\|_B$, где $\rho = \frac{1 - \sqrt{\eta}}{1 + 3\sqrt{\eta}}$, $\eta = \frac{\delta}{\Delta}$.

ПТИМ сходится на порядок быстрее метода простой итерации, метода Зейделя и метода Якоби. В практических задачах, когда m велико, отношение $\eta = \frac{\delta}{\Delta}$ часто является величиной порядка $O(m^{-2})$. Оценим скорость сходимости ПТИМ: $p_0(\varepsilon) = O(m)$. Оценим скорость сходимости метода простой итерации: $p_0(\varepsilon) = O(m^2)$.

10. Степенной метод решения частичной проблемы собственных значений

Рассмотрим частичную проблему собственных значений. Будем искать собственный вектор по формуле

$$x^{n+1} = Ax^n, \quad n \in \mathbb{Z}_+, \quad x^0 \text{ задано.} \tag{25}$$

Пусть $\{\lambda_k\}_{k=1}^m$ — собственные значения матрицы A , среди которых могут быть повторяющиеся. Упорядочим их по неубыванию модулей: $|\lambda_1| \leq |\lambda_2| \leq \dots \leq |\lambda_m|$. Будем доказывать сходимость степенного метода при выполнении трех условий:

- А) В вещественном пространстве $\mathbb{R}^m$ существует базис $\{e_k\}$, $k = \overline{1, m}$ из собственных векторов матрицы A .
- В) $\left| \frac{\lambda_1}{\lambda_2} \right| < 1$.
- С) $x^0 = c_1 e_1 + c_2 e_2 + \dots + c_m e_m$, где $c_m \neq 0$.

Утверждение. Пусть вещественная матрица A ($m \times m$) такова, что выполнены условия А) – С). Тогда степенной метод для матрицы A сходится по направлению к собственному вектору, отвечающему максимальному по модулю собственному значению: $x^n \xrightarrow{n \rightarrow \infty} e_m$. Кроме того, для последовательности $\left\{ \lambda_m^{(n)} \right\}$, заданной одной из формул $\lambda_m^{(n)} = \frac{x^{n+1}_i}{x^n_i}$, $i = \overline{1, m}$, либо $\lambda_m^{(n)} = \frac{(Ax^n, x^n)}{(x^n, x^n)}$ справедлива следующая оценка сходимости к λ_m : $\lambda_m^{(n)} - \lambda_m = O\left(\left(\frac{\lambda_{m-1}}{\lambda_m}\right)^n\right)$.

Замечание 3. Число итераций n , необходимое для вычисления максимального собственного значения с заданной точностью $\varepsilon > 0$, можно оценить из неравенства

$$\left(\frac{\lambda_{m-1}^{(n)}}{\lambda_m^{(n)}} \right)^n < \varepsilon.$$

Замечание. Пусть у вещественной матрицы A ($m \times m$) существует комплексное собственное значение: $\lambda = \lambda_0 + i\lambda_1$, $\lambda_1 \neq 0$. Тогда соответствующий собственный вектор — комплексный: $x = x_0 + ix_1$, $x_1 \neq \emptyset$, и начальное приближение x^0 вектора x в итерационном методе также должно быть комплексным.

11. Метод обратных итераций и обратных итераций со сдвигом

Метод обратных итераций

Пусть матрица A — невырожденная. Рассмотрим следующую форму записи неявного итерационного метода: $Ax^{n+1} = x^n$, $n \in \mathbb{Z}_+$, x^0 задано. Сформулируем три условия:

- А) В пространстве $\mathbb{R}^m$ существует базис $\{e_k\}$ из собственных векторов матрицы A .
- В) $\left| \frac{\lambda_1}{\lambda_2} \right| < 1$.
- С) $x^0 = c_1 e_1 + c_2 e_2 + \dots + c_m e_m$, $c_1 \neq 0$.

Утверждение. Пусть невырожденная вещественная матрица A ($m \times m$) такова, что выполнены условия А) – С). Тогда метод обратных итераций сходится по направлению к собственному вектору, отвечающему минимальному по модулю собственному значению: $x^n \xrightarrow{n \rightarrow \infty} e_1$.

Задача. Пусть выполнены условия А) – С) сходимости метода обратных итераций. Показать, что в случае произвольной матрицы A справедливы следующие оценки:

$$\lambda_1 - \frac{x^n_i}{x^{n+1}_i} = O\left(\left(\frac{\lambda_1}{\lambda_2}\right)^n\right), \quad \lambda_1 - \frac{(x^n, x^n)}{(x^{n+1}, x^n)} = O\left(\left(\frac{\lambda_1}{\lambda_2}\right)^n\right). \text{ Показать, что если матрица } A \text{ — самосопряженная, то последнюю оценку можно улучшить: } \lambda_1 - \frac{(x^n, x^n)}{(x^{n+1}, x^n)} = O\left(\left(\frac{\lambda_1}{\lambda_2}\right)^{2n}\right).$$

Метод обратных итераций со сдвигом

Рассмотрим итерационный метод, задаваемый формулой $(A - \alpha E)x^{n+1} = x^n$, $n \in \mathbb{Z}_+$, x^0 задано, где α — такое вещественное число, что матрица $(A - \alpha E)$ невырождена.

Само собственное значение λ_r находится из выражения: $\lambda_r = \lim_{n \rightarrow \infty} \left(\alpha + \frac{x^{(i)}_n}{x^{(i)}_{n+1}} \right)$, $i = \overline{1, m}$.

12. Приведение матрицы к к почти верхней треугольной форме при помощи элементарных отражений

Определение. Матрица A имеет верхнюю почти треугольную форму (ВПТФ), если ее можно записать в виде $A = \begin{pmatrix} \times & \times & \times & \dots & \times & \times \\ \times & \times & \times & \dots & \times & \times \\ 0 & \times & \times & \dots & \times & \times \\ 0 & 0 & \times & \dots & \times & \times \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \times & \times \end{pmatrix}$, где символами $\times$ обозначены, вообще говоря, ненулевые элементы матрицы.

Определение. Элементарным отражением, соответствующим вещественному вектор-столбцу $v = (v_1, v_2, \dots, v_m)^T$, называется преобразование, задаваемое матрицей

$$H = E - 2 \frac{vv^T}{\|v\|^2}. \tag{26}$$

$$vv^T = \begin{pmatrix} v_1^2 & v_1 v_2 & \dots & v_1 v_m \\ v_2 v_1 & v_2^2 & \dots & v_2 v_m \\ \vdots & \vdots & \ddots & \vdots \\ v_m v_1 & v_m v_2 & \dots & v_m^2 \end{pmatrix}$$

— симметричная (эрмитова) матрица.

Сформулируем свойства матрицы элементарного отражения:

1. H — симметрическая матрица, $H = H^T$.
2. H — ортогональная матрица, $H^{-1} = H^T$.

Утверждение. Пусть задан вещественный вектор-столбец $x = (x_1, x_2, \dots, x_m)^T$. Тогда можно выбрать вектор v так, чтобы было выполнено равенство $Hx = (-\|x\|, 0, \dots, 0)^T$, $\|x\| = \sqrt{(x, x)}$, где H — элементарное отражение, соответствующее вектор-столбцу v .

Утверждение. Любую вещественную матрицу A ($m \times m$) можно привести к верхней почти треугольной форме с помощью преобразования подобия с ортогональной матрицей Q : $C = Q^{-1}AQ =$

$$\begin{pmatrix} \times & \times & \times & \dots & \times & \times \\ \times & \times & \times & \dots & \times & \times \\ 0 & \times & \times & \dots & \times & \times \\ 0 & 0 & \times & \dots & \times & \times \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \times & \times \end{pmatrix}, \text{ где } Q^T = Q^{-1}.$$

Замечание 4. Преобразование подобия сохраняет спектр матрицы: $\lambda_k^C = \lambda_k^A$, $k = \overline{1, m}$.

Замечание 5. Если A — симметрическая матрица, то C также является симметрической матрицей: $A = A^T \Rightarrow C = C^T$.

Замечание 6. Симметричная матрица, имеющая верхнюю почти треугольную форму, является симметричной трездиагональной матрицей.

13. Понятие о QR алгоритме решения полной проблемы собственных значений

Утверждение. Произвольная матрица A ($m \times m$) может быть представлена в виде: $A = QR$, где Q — ортогональная матрица, а R — матрица, имеющая верхнюю треугольную форму (ВТФ).

Замечание. Число операций, необходимых для вычисления QR-разложения матрицы A , зависит от вида матрицы A . Для произвольной матрицы число операций можно оценить величиной порядка m^3 , для матрицы с ВПТФ, — порядка m^2 , для трездиагональной матрицы — порядка m .

Рассмотрим оптимальную версию QR-алгоритма. Приведем матрицу A к матрице A_0 , имеющей ВПТФ, и осуществим QR-разложение матрицы A_0 : $A_0 = Q_0R_0$, где Q_0 — ортогональная, а R_0 — верхнетреугольная матрица. Построим матрицу $A_1 = R_0Q_0$.

На следующем шаге осуществим QR-разложение матрицы $A_1 = Q_1R_1$ и построим матрицу $A_2 = R_1Q_1$. Аналогичным образом продолжая вычисления, на k -м шаге осуществим QR-разложение матрицы $A_k = Q_kR_k$ и построим $A_{k+1} = R_kQ_k$.

Утверждение. Если все собственные значения матрицы A вещественны, то последовательность матриц $\{A_k\}$ сходится к

матрице, имеющей ВТФ: $A_k \xrightarrow{k \rightarrow \infty} \begin{pmatrix} \lambda_1 & \times & \dots & \times \\ 0 & \lambda_2 & \dots & \times \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_m \end{pmatrix}$. Если

ли же матрица имеет комплексную пару собственных значений $\lambda_0 \pm i\lambda_1$, то ей на главной диагонали предельной матрицы будет соответствовать клетка размера 2×2 : $A_k \xrightarrow{k \rightarrow \infty}$

$\begin{pmatrix} \times & & & \times \\ & \times & & \\ & & \lambda_0 & \lambda_1 \\ & & -\lambda_1 & \lambda_0 \\ & & & \ddots \\ 0 & & & \times \end{pmatrix}$. (без доказательства)

Замечание 7. Итерационный процесс останавливается, когда все элементы ниже главной диагонали, либо ниже побочной (в случае комплексно-сопряженных собственных значений) матрицы A_n при некотором n становятся равными нулю. Однако следует заметить, что в данном случае под нулем мы понимаем либо машинный ноль, либо число, меньшее некоторой заданной величины — необходимой точности вычисления.

Замечание 8. QR-алгоритм применим к произвольной матрице A .

Замечание 9. QR-алгоритм является очень затратным по необходимому числу операций и объему памяти, используемому для хранения промежуточных матриц.

14. Метод простой итерации решения нелинейных уравнений. Сходимость метода

Рассмотрим функцию $f(x)$, $x \in \mathbb{R}$ и уравнение $f(x) = 0$. Будем считать, что начальное приближение $x^0 \in U_a(x^*)$ задано. Рассмотрим итерационные методы, задаваемые общей формулой

$$x^{n+1} = S(x^n), \quad n \in \mathbb{Z}_+$$

с некоторой функцией $S(x)$, определенной на $U_a(x^*)$. Пусть функция $S(x)$ имеет вид

$$S(x) = x + r(x)f(x), \quad S(x^*) = x^*, \quad (28)$$

где $r(x)$ — функция, не обращающаяся в ноль в окрестности $U_a(x^*)$, то есть $\operatorname{sgn}(r(x)) \neq 0$, $x \in U_a(x^*)$.

Определение. Итерационный метод, описываемый формулой (27) с функцией $S(x)$ вида (28), называется методом простой итерации.

Определение. Функция $S(x)$ удовлетворяет условию Липшица при $x \in U_a(x^*)$ с константой $q > 0$, если для любых точек $x_1, x_2 \in U_a(x^*)$ выполнено неравенство $|S(x_1) - S(x_2)| \leq q|x_1 - x_2|$.

Утверждение. Пусть функция $S(x)$ удовлетворяет условию Липшица с константой $q \in (0, 1)$ в некоторой окрестности $U_a(x^*)$, и пусть задано начальное приближение $x_0 \in U_a(x^*)$. Тогда метод простой итерации (27) сходится со скоростью геометрической прогрессии со знаменателем q .

Замечание. Если функция $S(x)$ непрерывно дифференцируема, то в качестве q можно взять максимальное значение $|S'(x)|$, и сходимость будет иметь место, если $q = \max_{x \in U_a(x^*)} |S'(x)| < 1$.

$$\frac{x^{n+1} - x^n}{\tau} + f(x^n) = 0, \quad \tau \in \mathbb{R}_+, \quad n \in \mathbb{Z}_+, \quad x^0 \in U_a(x^*). \quad (29)$$

Таким образом, если для поиска корня x^* применяется итерационный метод, записанный в виде (29) и $f'(x) > 0$, $x \in U_a(x^*)$, то значение параметра τ следует выбирать из интервала $(0, \frac{2}{M})$, где $M = \sup_{x \in U_a(x^*)} |f'(x)|$.

Метод Эйткена ускорения сходимости итерационного метода

Метод Эйткена позволяет ускорить сходимость метода простой итерации. Идея метода заключается в том, что после вычисления x^{n-1} , x^n , x^{n+1} производится пересчет по формуле $x'_{n+1} = x^{n+1} - \frac{(x^{n+1} - x^n)^2}{(x^{n+1} - 2x^n + x^{n-1})}$, и значение x'_{n+1} берется в качестве нового приближения.

15. Метод Ньютона решения нелинейных уравнений и систем нелинейных уравнений.

Рассмотрим функцию $f(x)$, $x \in \mathbb{R}$ и уравнение

$$f(x) = 0. \quad (30)$$

$$x^{n+1} = x^n - \frac{f(x^n)}{f'(x^n)}, \quad n \in \mathbb{Z}_+. \quad (31)$$

Определение. Итерационный процесс поиска корня уравнения (30), задаваемый формулой (31), называется итерационным методом Ньютона.

Замечание. Итерационный метод Ньютона часто называют методом касательных.

Замечание 10. Метод Ньютона является вычислительно сложным, поскольку на каждой итерации проводится вычисление значений производной функции $f(x)$, что является, вообще говоря, неустойчивым процессом.

Замечание 11. При решении задач на практике часто рассматривается модифицированный метод Ньютона, задаваемый формулой $x^{n+1} = x^n - \frac{f(x^n)}{f'(x^n)}$, $n \in \mathbb{Z}_+$. Преимущество этого метода перед классическим методом заключается в том, что в нем не требуется вычислять значения функции $f'(x)$ на каждой итерации. Однако при этом модифицированный метод Ньютона сходится медленнее классического метода Ньютона. Вопросы сходимости метода Ньютона излагаются в §4.

Метод Ньютона для нелинейных систем уравнений

$$\begin{cases} f_1(x_1, x_2) = 0 \\ f_2(x_1, x_2) = 0 \end{cases}.$$

Введем векторы $f = (f_1, f_2)^T$, $x = (x_1, x_2)^T$ и матрицу Якоби системы (32) — матрицу из частных производных функций $f_1(x)$ и $f_2(x)$:

$$J(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \frac{\partial f_1}{\partial x_2}(x) \\ \frac{\partial f_2}{\partial x_1}(x) & \frac{\partial f_2}{\partial x_2}(x) \end{pmatrix}. \quad (33)$$

Замечание. При поиске значения каждой следующей итерации x^{n+1} можно сначала решить линейную систему: $J(x^n)v^n = -f(x^n)$, $n \in \mathbb{Z}_+$, где $v^n = x^{n+1} - x^n$. Теперь значение x^{n+1} получается из найденного v^n : $x^{n+1} = x^n + v^n$.

$$\begin{cases} f_1(x_1, x_2, \dots, x_m) = 0 \\ f_2(x_1, x_2, \dots, x_m) = 0 \\ \dots \\ f_m(x_1, x_2, \dots, x_m) = 0 \end{cases}.$$

Введем векторы $f = (f_1, f_2, \dots, f_m)^T$, $x = (x_1, x_2, \dots, x_m)^T$ и матрицу Якоби системы (34): $J = (f_{ij})$, $f_{ij} = \frac{\partial f_i}{\partial x_j}$, $i, j = \overline{1, m}$. Запишем схему итерационного метода Ньютона, используя матрицу Якоби: $x^{n+1} = x^n - J^{-1}(x^n)f(x^n)$, $n \in \mathbb{Z}_+$.

Заметим, что вычислять матрицу J на каждом шаге достаточно трудно.

Замечание. Аналогично одномерному случаю можно рассматривать модифицированный метод Ньютона для решения нелинейных систем: $x^{n+1} = x^n - J^{-1}(x^n)f(x^n)$, $n \in \mathbb{Z}_+$. Реализация модифицированного метода Ньютона проще классического варианта, но скорость сходимости при данном подходе меньше.

16. Сходимость метода Ньютона. Оценка скорости сходимости

Теорема 7. Пусть существует такая константа $M > 0$, для которой выполнена оценка $\frac{1}{2}|S''(x)| \leq M$, $x \in U_a(x^*)$. Тогда если начальное приближение x^0 выбрать в соответствии с условием $|x^0 - x^*| < \frac{1}{M}$, то итерационный метод Ньютона сходится, и имеет место оценка: $|x^n - x^*| \leq \frac{1}{M}(M|x^0 - x^*|)^{2^n}$.

Замечание 12. Если итерационный метод Ньютона сходится, то достаточно быстро. При наличии оценки вида

$$|z^{n+1}| \leq M|z^n|^2,$$

говорят о квадратичной сходимости метода.

Замечание 13. Из условий теоремы следует, что начальное приближение нужно выбирать достаточно близко к точному решению рассматриваемого уравнения.

Замечание 14. Другие рассмотренные нами методы (модифицированный метод Ньютона и метод секущих) обладают, по крайней мере, линейной скоростью. Это следует из того, что если из записать в виде $x^{n+1} = S(x^n)$, то $S(x^*) = x^*$ и $S'(x^*) \neq 0$. Например, для модифицированного метода Ньютона $S'(x^*) = 1 - \frac{f'(x^*)}{f'(x^0)}$, и чем ближе взять x^0 к x^* , тем быстрее будет сходимость.

17. Явная разностная схема. Погрешность, сходимость, устойчивость.

$$\frac{\partial u(x,t)}{\partial t}=\frac{\partial^2 u(x,t)}{\partial x^2}+f(x,t),\quad (x,t)\in G=\{(x,t):\;x\in(0,1),\;t\in(0,T)\},\quad (36)$$

$$\begin{cases} u(0,t)=\mu_1(t)\\ u(1,t)=\mu_2(t), \end{cases}\quad t\in[0,T],\quad (37)$$

$$u(x,0)=u_0(x),\quad x\in[0,1]\quad (38)$$

Определение. *Сеточной функцией называется функция дискретного аргумента на заданной сетке, то есть такая функция определена только в узлах данной сетки.*

$$\frac{y_i^{n+1}-y_i^n}{\tau}=\frac{y_{i-1}^n-2y_i^n+y_{i+1}^n}{h^2}+f(x_i,t_n),\quad (x_i,t_n)\in\omega_{\tau h}.\quad (39)$$

$$\begin{cases} y_0^{n+1}=\mu_1(t_{n+1})\\ y_N^{n+1}=\mu_2(t_{n+1}), \end{cases}\quad t_{n+1}\in\bar\omega_\tau,\quad (40)$$

$$y_i^0=u_0(x_i),\quad x_i\in\bar\omega_h,\quad y_i^0=u_0(x_i),\quad x_i\in\bar\omega_h.\quad (41)$$

Определение. *Дискретным аналогом задачи (36)–(38), или ее разностной схемой, называется система линейных уравнений (39)–(41).*

Замечание 15. *В первой краевой задаче численные значения решения y_0^{n+1} и y_N^{n+1} равны значениям функций $\mu_1(t)$ и $\mu_2(t)$ соответственно при $t=t_{n+1}$ (хотя это и не обязательно). В случае краевых условий иного типа, аппроксимация краевых условий должна быть согласована по порядку погрешности с порядком аппроксимации уравнения. Определение аппроксимации и порядка погрешности аппроксимации будет дано ниже.*

Замечание 16. *Заметим, что в уравнении (39) значения функции $f(x,t)$ не обязательно брать именно в узлах рассматриваемой сетки, можно использовать значения этой функции с некоторой «поправкой». Что именно имеется в виду под «поправкой», будет рассмотрено далее, а также будет показано, что выбор значений функции $f(x,t)$ для разностной схемы, использующих такую «поправку», позволит получить более высокий порядок погрешности аппроксимации, а стало быть и более точное решение исходного уравнения.*

Замечание 17. *Качество и скорость решения численной задачи (39)–(41) во многом зависят от выбора числа узлов сетки $\omega_{\tau h}$: чем меньше узлов в сетке, тем меньше уравнений содержится в системе, тем проще и быстрее ее решать, но и приближение решения исходной задачи в этом случае будет более грубым.*

Замечание. *Вопросы сходимости и устойчивости разностной схемы являются ключевыми, однако обычно достаточно рассмотреть только один из этих двух вопросов: в конце курса будет доказано, что из устойчивости разностной схемы следует ее сходимость к решению исходной задачи при условии, что разностная схема аппроксимирует исходную задачу.*

Определение. *Совокупность узлов, которые участвуют в записи разностной схемы, называют шаблоном.*

Определение. *Сеточная функция вида*

$$z_i^n=z(x_i,t_n)=y_i^n-u_i^n,\quad (x_i,t_n)\in\bar\omega_{\tau h}\quad (42)$$

называется погрешностью решения разностной схемы (39)–(41).

$$\psi_i^n=\frac{u_{i-1}^n-2u_i^n+u_{i+1}^n}{h^2}-\frac{u_i^{n+1}-u_i^n}{\tau}+f_i^n,\quad (43)$$

Определение. *Сеточная функция, задаваемая равенствам (43) называется погрешностью аппроксимации разностной схемы (39)–(41) на решении исходной задачи.*

Задача. *Доказать, что $\psi_i^n=O(\tau+h^2)$.*

Теорема. *Пусть решение $u(x,t)$ задачи (36)–(38) обладает достаточной гладкостью (четыре раза дифференцируема по x и два раза по t). Тогда для сходимости решения разностной схемы (39)–(41) к решению исходной задачи (36)–(38) в норме $\|\cdot\|_C$ необходимо и достаточно, чтобы выполнялось условие:*

$$\gamma=\frac{\tau}{h^2}\leqslant 0.5.$$

При этом условии, выполняется оценка: $\|z^{n+1}-u^{n+1}\|_C\leqslant M_1(\tau+h^2)$, $n=0,1,\ldots$ где $M_1>0$ — константа, не зависящая от τ и h .

Замечание 18. *Разностные схемы могут сходиться условно (и быть условно устойчивыми) и абсолютно. Условная сходимость определяется наличием ограничений на шаги сетки любого характера, для абсолютной сходимости требуется, чтобы такие ограничения отсутствовали. В примененной выше теореме условие сходимости имеет вид $\frac{\tau}{h^2}\leqslant 0.5$. Следовательно, явная разностная схема является условно сходящейся.*

Замечание 19. *Важно помнить, что сходимость и устойчивость разностной схемы доказывается в конкретной норме. В данном параграфе доказана сходимость и устойчивость решений разностной схемы (39)–(41) в норме $\|\cdot\|_C$, которая является достаточно сильной нормой, а значит, обеспечивает более точную оценку, по сравнению, например, со среднеквадратичной нормой.*

18. Чисто неявная разностная схема. Аппроксимация, сходимость, устойчивость

$$\frac{\partial u(x,t)}{\partial t}=\frac{\partial^2 u(x,t)}{\partial x^2}+f(x,t),\quad (44)$$

$$(x,t)\in G=\{(x,t):\;x\in(0,1),\;t\in(0,T)\}$$

$$\begin{cases} u(0,t)=\mu_1(t)\\ u(1,t)=\mu_2(t), \end{cases}\quad t\in[0,T],u(x,0)=u_0(x),\quad x\in[0,1].\quad (45)$$

Поставим в соответствие задаче (44)–(45) следующую разностную схему:

$$\frac{y_i^{n+1}-y_i^n}{\tau}=\frac{y_{i-1}^{n+1}-2y_i^{n+1}+y_{i+1}^{n+1}}{h^2}+f(x_i,t_{n+1}),\quad (46)$$

$$(x_i,t_n),(x_i,t_{n+1})\in\omega_{\tau h},$$

$$\begin{cases} y_0^{n+1}=\mu_1(t_{n+1})\\ y_N^{n+1}=\mu_2(t_{n+1}), \end{cases}\quad t_{n+1}\in\bar\omega_\tau,y_i^0=u_0(x_i),\quad x_i\in\bar\omega_h,\quad (47)$$

где $y_i^n=y(x_i,t_n)$ —искомое численное решение в точке $(x_i,t_n)\in\bar\omega_{\tau h}$.

Эта система имеет трехдиагональную матрицу порядка $(N-1)$:

$$A=\begin{pmatrix} 1+2\gamma & -\gamma & 0 & \ldots & 0 & 0 \\ -\gamma & 1+2\gamma & -\gamma & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1+2\gamma & -\gamma \\ 0 & 0 & 0 & \ldots & -\gamma & 1+2\gamma \end{pmatrix},$$

обладающую строгим диагональным преобладанием: $a_{ii}>\sum_{\substack{j=1\\j\neq i}}^N|a_{ij}|$, $i=\overline{1,(N-1)}$.

ψ_i^n — погрешность аппроксимации на решении:

$$\psi_i^n=\psi(x_i,t_n)=-\frac{u_i^{n+1}-u_i^n}{\tau}+\frac{u_{i-1}^{n+1}-2u_i^{n+1}+u_{i+1}^{n+1}}{h^2}+f(x_i,t_{n+1}).\quad (48)$$

Задача. *Доказать, что $\psi_i^n=O(\tau+h^2)$.*

Теорема. *Пусть функция $u(x,t)$ имеет достаточную гладкость (четыре раза дифференцируема по x и два раза по t). Тогда чисто неявная разностная схема сходится к решению исходной задачи в норме $\|\cdot\|_C$ с первым порядком точности по τ и вторым порядком точности по h .*

Замечание. *Если в разностной задаче (46)–(47) взять нулевые краевые условия $y_0^{n+1}=y_N^{n+1}=0$, то для y_i^n можно вывести оценку, аналогичную полученной выше: $\|y^{n+1}\|_C\leqslant\|u_0\|_C+\tau\sum_{k=0}^N\|f^k\|_C$. Эта оценка означает, что решение разностной схемы устойчиво по начальному условию и по правой части уравнения.*

19. Симметричная разностная схема. Аппроксимация, сходимость, устойчивость

Рассмотрим уравнение теплопроводности с краевыми условиями первого рода:

$$\frac{\partial u(x,t)}{\partial t}=\frac{\partial^2 u(x,t)}{\partial x^2}+f(x,t),\quad (49)$$

$$(x,t)\in G=\{(x,t):\;x\in(0,1),\;t\in(0,T]\},\quad (50)$$

$$\begin{cases} u(0,t)=\mu_1(t)\\ u(1,t)=\mu_2(t), \end{cases}\quad t\in[0,T],\quad (51)$$

$$u(x,0)=u_0(x),\quad x\in[0,1].\quad (52)$$

Воспользуемся сетками $\omega_{\tau h}$ и $\bar\omega_{\tau h}$, введенными в первом параграфе данной главы на множествах G и $\bar G$ соответственно.

Введем вторую разностную производную для дискретной функции $y_i^n=y(x_i,t_n)$, определенной на множестве $\bar\omega_{\tau h}$:

$$y_{\bar{x}x,i}^n=\frac{y_{i-1}^n-2y_i^n+y_{i+1}^n}{h^2}.$$

Эта производная является дискретным аналогом второй производной по x функции $u(x,t)$.

Поставим в соответствие уравнению (49) его дискретный аналог в виде

$$\frac{y_i^{n+1}-y_i^n}{\tau}=\frac{y_{\bar{x}x,i}^{n+1}+y_{\bar{x}x,i}^n}{2}+f(x_i,t_{n+\frac{1}{2}}),\quad (53)$$

где $(x_i,t_{n+\frac{1}{2}})=(x_i,t_n+\frac{\tau}{2})\in\omega_{\tau h}$.

Определение. *Слой $t_{n+\frac{1}{2}}=t_n+\frac{\tau}{2}$ называется полуцелым слоем.*

Добавим краевые и начальные условия:

$$\begin{cases} y_0^{n+1}=\mu_1(t_{n+1})\\ y_N^{n+1}=\mu_2(t_{n+1}), \end{cases}\quad t_{n+1}\in\bar\omega_\tau,\quad (54)$$

$$y_i^0=u_0(x_i),\quad x_i\in\bar\omega_h.\quad (55)$$

В рассматриваемой разностной схеме использован шеститочечный шаблон вида

$$\begin{array}{ccccccc} & \times & & \times & & \times & \\ & i-1 & & i & & i+1 & \\ & \times & & \times & & \times & \\ & i-1 & & i & & i+1 & \end{array}\quad \begin{array}{c} t_{n+1} \\ \\ t_n \end{array}$$

Заметим, что данная схема, с точки зрения нахождения численного решения, похожа на ту, которую мы рассматривали в предыдущем параграфе, в частности, матрица системы, соответствующей этой схеме, является трехдиагональной со строгим диагональным преобладанием. Это значит, что решение разностной схемы (53)–(55) такой задачи существует, единственно и находится с помощью метода прогонки.

Введем погрешность решения разностной схемы:

$$z_i^n=z(x_i,t_n)=y_i^n-u_i^n,$$

где $u_i^n=u(x_i,t_n)$, $(x_i,t_n)\in\bar\omega_{\tau h}$.

Выразив y_i^n из этого выражения и подставив его в уравнение (53), получим задачу относительно z_i^n :

$$\frac{z_i^{n+1}-z_i^n}{\tau}=\frac{z_{\bar{x}x,i}^{n+1}+z_{\bar{x}x,i}^n}{2}+\psi_i^n,\quad (x_i,t_n),(x_i,t_{n+1})\in\omega_{\tau h},\quad (56)$$

$$\begin{cases} z_0^{n+1}=0\\ z_N^{n+1}=0, \end{cases}\quad t_{n+1}\in\bar\omega_\tau,\quad (57)$$

$$z_i^0=0,\quad x_i\in\bar\omega_h,\quad (58)$$

где ψ_i^n — погрешность аппроксимации на решении исходной задачи (49)–(52):

$$\psi_i^n=\psi(x_i,t_n)=-\frac{u_i^{n+1}-u_i^n}{\tau}+\frac{u_{\bar{x}x,i}^{n+1}+u_{\bar{x}x,i}^n}{2}+f(x_i,t_{n+\frac{1}{2}}),\quad (59)$$

$$(x_i,t_n),(x_i,t_{n+1})\in\omega_{\tau h}.\quad (60)$$

Решение Разложим u_i^{n+1} и u_i^n в ряд Тейлора в окрестности точки $(x_i,t_{n+\frac{1}{2}})$:

$$u_i^{n+1}=u_i^{n+\frac{1}{2}}+u_{t,i}^{n+\frac{1}{2}}\frac{n+\frac{1}{2}}{2}+\frac{1}{2}u_{tt,i}^{n+\frac{1}{2}}\left(\frac{\tau}{2}\right)^2+O(\tau^3),$$

$$u_i^n=u_i^{n+\frac{1}{2}}-u_{t,i}^{n+\frac{1}{2}}\frac{n+\frac{1}{2}}{2}+\frac{1}{2}u_{tt,i}^{n+\frac{1}{2}}\left(\frac{\tau}{2}\right)^2+O(\tau^3).$$

Подставим эти разложения в формулу для ψ_i^n :

$$\begin{aligned} \psi_i^n &= -\frac{u_i^{n+1}-u_i^n}{\tau}+\frac{u_{\bar{x}x,i}^{n+1}+u_{\bar{x}x,i}^n}{2}+f(x_i,t_n+0.5\tau)= \\ &= -u_{t,i}^{n+\frac{1}{2}}+O(\tau^2)+\frac{u_{\bar{x}x,i}^{n+1}+u_{\bar{x}x,i}^n}{2}+f_i^{n+\frac{1}{2}}. \end{aligned}$$

Теперь в представлении второй разностной производной разложим все вхождения функции $u(x_i,t_n)$ в ряд Тейлора по переменной x в окрестности точки $(x_i,t_{n+\frac{1}{2}})$. Приведа подобные слагаемые, получим

$$u_{\bar{x}x,i}^n=u_{\bar{x}x,i}^n+u_{x^4,i}^n\frac{h^2}{12}+O(h^4).$$

Применим это разложение для $u_{\bar{x}x,i}^{n+1}$ затем проведем еще одно разложение в ряд Тейлора в точке $(x_i,t_{n+\frac{1}{2}})$, теперь по переменной t :

$$\begin{aligned} u_{\bar{x}x,i}^{n+1} &= u_{\bar{x}x,i}^{n+1}+u_{\bar{x}x,i}^{n+1}\frac{h^2}{12}+O(h^4)= \\ &= u_{\bar{x}x,i}^{n+\frac{1}{2}}+u_{\bar{x}x,t,i}\frac{n+\frac{1}{2}}{2}+u_{x^4,i}^{n+1}\frac{h^2}{12}+u_{\bar{x}x,i}^{n+1}\frac{h^2}{12}\frac{\tau}{2}+O(\tau^2+h^4). \end{aligned}$$

То же самое проделаем для $u_{\bar{x}x,i}^n$:

$$\begin{aligned} u_{\bar{x}x,i}^{n+1} &= u_{\bar{x}x,i}^{n+1}+u_{x^4,i}^{n+1}\frac{h^2}{12}+O(h^4)= \\ &= u_{\bar{x}x,i}^{n+\frac{1}{2}}-u_{\bar{x}x,t,i}\frac{n+\frac{1}{2}}{2}+u_{x^4,i}^{n+1}\frac{h^2}{12}-u_{\bar{x}x,t,i}^{n+1}\frac{h^2}{12}\frac{\tau}{2}+O(\tau^2+h^4). \end{aligned}$$

Подставим эти разложения в формулу (59) для ψ_i^n и учтем выполнение уравнения теплопроводности (49):

$$\psi_i^n=(-u_{t,i}^{n+\frac{1}{2}}+u_{\bar{x}x,i}^{n+\frac{1}{2}}+f_i^{n+\frac{1}{2}})+u_{x^4,i}^{n+\frac{1}{2}}\frac{h^2}{12}+O(\tau^2+h^4)=O(\tau^2+h^2).$$

Теорема. *Пусть функция $u(x,t)$, являющаяся решением задачи для уравнения теплопроводности (49)–(52), имеет достаточную гладкость. Тогда симметричная разностная схема (53)–(55) сходится к решению исходной задачи со вторым порядком по τ и вторым порядком по h в $L_2(\omega_h)$ -норме пространства сеточных функций.*

Замечание. *Если в разностной задаче (53)–(55) взять нулевые краевые условия*

$$y_0^{n+1}=y_N^{n+1}=0,$$

то для y_i^n можно ввести априорную оценку, аналогичную полученной выше оценке (??):

$$\|y^{n+1}\|_{L_2(\omega_h)}\leqslant\|u_0\|_{L_2(\omega_h)}+\tau\sum_{j=0}^n\|f(t_j)\|_{L_2(\omega_h)}.$$

Эта оценка означает, что решение разностной схемы устойчиво в норме $L_2(\omega_h)$ по начальному условию u_0 и правой части уравнения.

20. Разрешимость разностной задачи. Сходимость разностной задачи Дирихле

Продолжаем рассматривать задачу Дирихле

$$\begin{cases} \frac{\partial^2 u}{\partial x_1^2}+\frac{\partial^2 u}{\partial x_2^2}=f(x_1,x_2),\;(x_1,x_2)\in G,\\ u(x_1,x_2)|_{\Gamma}=\mu(x_1,x_2) \end{cases}\quad (61)$$

Запишем разностную схему (??) из ?? в виде:

$$\begin{cases} \frac{y_{i-1,j}-2y_{ij}+y_{i+1,j}}{h_1^2}+\frac{y_{i,j-1}-2y_{ij}+y_{i,j+1}}{h_2^2}=f_{ij},\\ i=\overline{1,(N_1-1)},\;j=\overline{1,(N_2-1)},\\ y_{ij}|_{\Gamma_h}=\mu_{ij}. \end{cases}$$

Напомним, что f_{ij},μ_{ij} — значения непрерывных функций $f(x_1,x_2)$ и $\mu(x_1,x_2)$ в узлах сетки ω_h . Разрешим эту схему относительно центрального узла x_{ij} :

$$\begin{cases} \left(\frac{2}{h_1^2}+\frac{2}{h_2^2}\right)y_{ij}=\frac{y_{i-1,j}+y_{i+1,j}}{h_1^2}+\frac{y_{i,j-1}+y_{i,j+1}}{h_2^2}-f_{ij},\\ i=\overline{1,(N_1-1)},\;j=\overline{1,(N_2-1)},\\ y_{ij}|_{\Gamma_h}=\mu_{ij}. \end{cases}\quad (62)$$

Для того чтобы эта система имела решение при любых значениях функций $f(x_1,x_2)$ и $\mu(x_1,x_2)$, необходимо и достаточно, чтобы однородная система линейных уравнений имела только тривиальное решение.

Пусть H_{N_1-1,N_2-1} — пространство сеточных функций, определенных на сетке ω_h и обращающихся в нуль на границе Γ_h . Введем норму в этом пространстве:

$$\|v\|_C=\max_{\substack{1\leqslant i\leqslant N_1-1\\1\leqslant j\leqslant N_2-1}}|v_{ij}|,\;v\in H_{N_1-1,N_2-1}.$$

Теорема 8. *Однородная система линейных уравнений*

$$\begin{cases} \left(\frac{2}{h_1^2}+\frac{2}{h_2^2}\right)v_{ij}=\frac{v_{i-1,j}+v_{i+1,j}}{h_1^2}+\frac{v_{i,j-1}+v_{i,j+1}}{h_2^2},\\ i=\overline{1,(N_1-1)},\;j=\overline{1,(N_2-1)},\\ v_{ij}|_{\Gamma_h}=0 \end{cases}$$

имеет единственное решение, и оно является тривиальным:

$$v_{ij}=0,\;x_{ij}\in\bar\omega_h.$$

Следствие. *Разностная задача*

$$\begin{cases} y_{x_1x_1,ij}+y_{x_2x_2,ij}=f_{ij},\quad x_{ij}=\left(x_1^{(i)},x_2^{(j)}\right)\in\omega_h,\\ y_{ij}|_{\Gamma_h}=\mu_{ij} \end{cases}$$

имеет единственное решение при любых значениях f_{ij} и μ_{ij} , $x_{ij}\in\omega_h$.

Лемма (принцип максимума). *Пусть для сеточной функции v , определенной на сетке ω_h , выполнены неравенства*

$$v_{ij}\geqslant 0,\;x_{ij}\in\Gamma_h,$$

$$L_h v_{ij}\geqslant 0,\;x_{ij}\in\omega_h.$$

Тогда справедливо следующее неравенство:

$$v_{ij}\geqslant 0,\;x_{ij}\in\bar\omega_h.$$

Следствие. *Рассмотрим две разностные задачи*

$$L_h y_{ij}=\varphi_{ij},\;x_{ij}\in\omega_h,\;y_{ij}|_{\Gamma_h}-\text{ заданы,}$$

$$L_h Y_{ij}=\Phi_{ij},\;x_{ij}\in\omega_h,\;Y_{ij}|_{\Gamma_h}-\text{ заданы.}$$

Если выполнены неравенства

$$|y_{ij}|\leqslant Y_{ij},\;x_{ij}\in\Gamma_h,$$

$$|\varphi_{ij}|\leqslant\Phi_{ij},\;x_{ij}\in\omega_h,$$

то справедливо следующее неравенство:

$$|y_{ij}|\leqslant Y_{ij},\;x_{ij}\in\bar\omega_h.$$

Теорема 9. *Пусть решение задачи (61) четыре раза непрерывно дифференцируемо в $\bar G$. Тогда решение разностной задачи (??) сходится к решению исходной задачи со сеточной норме C , и имеет место оценка*

$$\|y_{ij}-u\left(x_1^{(i)},x_2^{(j)}\right)\|_C\leqslant M\left(h_1^2+h_2^2\right),$$

где $M>0$ — константа, не зависящая от h_1 и h_2 .

21. Методы решения разностной задачи Дирихле

Рассмотрим разностную задачу Дирихле:

$$\begin{cases} \left(\frac{2}{h_1^2}+\frac{2}{h_2^2}\right)y_{ij}=\frac{y_{i-1,j}+y_{i+1,j}}{h_1^2}+\frac{y_{i,j-1}+y_{i,j+1}}{h_2^2}-f_{ij},\\ i=\overline{1,(N_1-1)},\;j=\overline{1,(N_2-1)},\\ y_{ij}|_{\Gamma_h}=\mu_{ij}. \end{cases}\quad (63)$$

Для нахождения решения этой разностной схемы нужно решить СЛАУ с матрицей порядка $(N_1-1)\times (N_2-1)$. Заметим, что матрица системы разрежена, то есть среди элементов этой матрицы содержится большое число нулей. Очевидно, что использование классического метода Гаусса для решения такой системы не будет оптимальным. Существуют значительно более эффективные как прямые, так и итерационные методы решения системы (63) (см. [7]). Рассмотрим несколько итерационных методов, решающих поставленную задачу: методы Якоби, Зейделя и попеременно-треугольный итерационный метод.

Метод Якоби

Итерационный процесс задается схемой

$$\begin{cases} \left(\frac{2}{h_1^2}+\frac{2}{h_2^2}\right)y_{ij}^{(s+1)}=\frac{y_{i-1,j}^{(s)}+y_{i+1,j}^{(s)}}{h_1^2}+\frac{y_{i,j-1}^{(s)}+y_{i,j+1}^{(s)}}{h_2^2}-f_{ij},\\ i=\overline{1,(N_1-1)},\;j=\overline{1,(N_2-1)},\\ y_{ij}^{(s+1)}|_{\Gamma_h}=\mu_{ij}, \end{cases}$$

где $s\in\mathbb{Z}_+$, $y_{ij}^{(0)}$ — задано.

Метод Зейделя

В методе Зейделя итерации определяются по правилу:

$$\begin{cases} \left(\frac{2}{h_1^2}+\frac{2}{h_2^2}\right)y_{ij}^{(s+1)}=\\ \frac{1}{h_1^2}\left(y_{i-1,j}^{(s+1)}+y_{i+1,j}^{(s)}\right)+\frac{1}{h_2^2}\left(y_{i,j+1}^{(s+1)}+y_{i,j-1}^{(s)}\right)-f_{ij},\\ i=\overline{1,N_1-1},\;j=\overline{1,N_2-1},\;s=0,1,2,\ldots,\\ y_{ij}^{(s+1)}|_{\Gamma_h}=\mu_{ij}. \end{cases}$$

Несмотря на то, что метод Зейделя формально является неявным, нетрудно предложить алгоритм вычисления решения на $s+1$ -й итерации по явным формулам. Порядок счета здесь следующий. Сначала находятся значения $y_{1,j}^{(s+1)}$, $j=\overline{1,N_2-1}$, используя $y_{ij}^{(s)}$ и граничные значения μ_{ij} . Затем, используя $y_{1,j}^{(s+1)}$, находят $y_{2,j}^{(s+1)}$, $j=\overline{1,N_2-1}$ и т.д. Это означает, что вычисление значений $y_{i,j}^{(s+1)}$ на новой итерации осуществляется последовательно от левого нижнего угла (точки $x_{1,1}$) до правого верхнего угла (точки x_{N_1-1,N_2-1}).

Попеременно-треугольный итерационный метод

Эффективным методом решения системы разностных уравнений (63) является попеременно-треугольный итерационный метод. Запишем систему (63) в матричной форме

$$Ay=\varphi$$

с симметричной положительно определенной матрицей A и представим матрицу A в виде суммы

$$A=R_1+R_2,$$

где R_1 — нижняя треугольная, а R_2 — верхняя треугольная матрица. На главных диагоналях R_1 и R_2 стоят элементы $0.5a_{ii}$, где a_{ii} — элементы A , стоящие на главной диагонали. Попеременно-треугольный итерационный метод имеет вид

$$(E+\omega R_1)(E+\omega R_2)\frac{y^{(s+1)}-y^{(s)}}{\tau}+Ay^{(s)}=\varphi,$$

где E — единичная матрица, ω и τ — числовые параметры. Метод сходится при $\omega>\frac{\tau}{4}>0$. В случае системы (63) матрицы R_1 и R_2 определяются соотношениями

$$(R_1y)_{ij}=\frac{y_{ij}-y_{i-1,j}}{h_1^2}+\frac{y_{ij}-y_{i,j-1}}{h_2^2},$$

$$(R_2y)_{ij}=\frac{y_{ij}-y_{i+1,j}}{h_1^2}+\frac{y_{ij}-y_{i,j+1}}{h_2^2}.$$

Алгоритм нахождения $y^{(s+1)}$ сводится к последовательному решению двух уравнений

$$(E+\omega R_1)W^{(s)}=\varphi-Ay^{(s)},$$

$$(E+\omega R_2)\frac{y^{(s+1)}-y^{(s)}}{\tau}=W^{(s)},$$

каждое из которых решается путем обращения треугольных матриц.

Замечание. Показано, что попеременно-треугольный итерационный метод решения систем линейных алгебраических уравнений требует для достижения заданной точности $\epsilon>0$, числа итераций $\mathbf{n_0}(\epsilon)$ на порядок меньше, чем методы Якоби, Зейделя, простой итерации. В силу этого он широко применяется для решения практических задач.

22. Основные понятия теории разностных схем: аппроксимация, устойчивость, сходимость.

Пусть дана исходная дифференциальная задача. Не конкретизируя вид этой задачи, запишем ее в форме операторного уравнения.

$$L(u(x))=f(x),\quad x\in G, \quad (64)$$

где G — область изменения независимых переменных x (аргумент x может быть многомерным), $f(x)$ — заданная функция, L — линейный дифференциальный оператор. Предполагается, что начальные и граничные условия учитываются видом оператора L и правой частью $f(x)$. Будем считать, что исходная задача корректно поставлена. Это означает, что ее решение существует, оно единственно и непрерывно зависит от правой части $f(x)$.

Для построения разностной схемы, прежде всего в области G вводится разностная сетка G_h — конечное множество точек, принадлежащих G . Точки $x\in G_h$ называются узлами сетки. Параметр h (шаг сетки) характеризует плотность заполнения области G точками G_h . Будем считать h вектором, для которого определена норма $|h|$. Фактически имеют дело с последовательностью сеток, число узлов которых $N=N(h)$ увеличивается с уменьшением $|h|$. Обычно число узлов $N=N(h)$ сетки G_h неограниченно возрастает при $|h|\rightarrow 0$. Конкретные примеры разностных сеток были приведены в предыдущих параграфах.

После того, как введена сетка G_h , функции непрерывного аргумента x , определенные для $x\in G$, заменют сеточными функциями, т.е. функциями, определенными только в точках сетки G_h . Правую часть $f(x)$ уравнения (64) заменяют приближенно некоторой сеточной функцией $\varphi_h(x)$, $x\in G_h$, а дифференциальный оператор L — линейным разностным оператором L_h . В результате вместо дифференциального уравнения (64) получают систему разностных уравнений

$$L_hy_h(x)=\varphi_h(x),\quad x\in G_h, \quad (65)$$

которая называется разностной схемой. Для изучения устойчивости и сходимости разностной задачи (65) нужно ввести пространство сеточных функций. Будем считать, что решение $u(x)$ задачи (64) принадлежит линейному нормированному пространству B_0 , а решение $y_h(x)$ разностной схемы (65) принадлежит конечномерному линейному нормированному пространству B_h . Нормы в пространствах B_0 и B_h будем обозначать, соответственно, через $\|\cdot\|_0$ и $\|\cdot\|_h$.

Существенным при введении конкретных норм является вопрос о связи этих норм в пространствах B_0 и B_h .

Введем оператор проектирования $P_h:B_0\rightarrow B_h$, т.е. линейный оператор, сопоставляющий каждому элементу $u\in B_0$ некоторый элемент $u_h\in B_h$. Будем обозначать через u_h проекцию элемента u

$$P_hu=u_h,\;u\in B_0,\;u_h\in B_h. \quad (66)$$

Будем предполагать в дальнейшем, что нормы $\|\cdot\|_0$ и $\|\cdot\|_h$ согласованы в том смысле, что для любого элемента $u\in B_0$ существует

$$\lim_{|h|\rightarrow 0}\|u_h\|_h=\|u\|_0. \quad (67)$$

Определение. *Сеточная функция*

$$z_h(x)=y_h(x)-u_h(x),\;x\in G_h,$$

называется *погрешностью разностной схемы* (65).

Подставим выражение $y_h(x)=z_h(x)+u_h(x)$ в уравнение (65) и, используя линейность оператора L_h , получим

$$L_hz_h(x)+L_hu_h(x)=\varphi_h(x).$$

Отсюда следует

$$L_hz_h(x)=\psi_h(x),\;x\in G_h, \quad (68)$$

где

$$\psi_h(x)=\varphi_h(x)-L_hu_h(x). \quad (69)$$

Определение. *Сеточная функция $\psi_h(x)$, определенная по формуле (69), называется погрешностью аппроксимации (или невязкой) разностной задачи (65) на решении исходной дифференциальной задачи (64).*

Таким образом, для погрешности разностной схемы $z_h(x)$ получаем задачу (68) той же структуры, что и разностная задача (65), но с правой частью, представляющей собой погрешность аппроксимации на решении.

Спроектируем исходное уравнение (64) на пространство B_h , т.е. запишем сеточное уравнение

$$P_h(Lu)=P_hf.$$

Представим погрешность аппроксимации $\psi_h(x)$ в виде суммы

$$\psi_h(x)=\psi_h^{(1)}(x)+\psi_h^{(2)}(x),$$

где $\psi_h^{(1)}(x)=P_h(Lu)(x)-L_hu_h(x)$, $\psi_h^{(2)}(x)=\varphi_h(x)-f_h(x)$, $f_h(x)=P_hf$. Функция $\psi_h^{(1)}(x)$ называется погрешностью аппроксимации дифференциального оператора L разностным оператором L_h . Функция $\psi_h^{(2)}(x)$ называется погрешностью аппроксимации правой части.

Определение. *Говорят, что разностная схема (65) аппроксимирует исходную задачу (64), если $\|\psi_h\|_h\rightarrow 0$ при $|h|\rightarrow 0$.*

Определение. *Разностная схема (65) имеет k-й порядок аппроксимации, если существуют положительные константы M_1 и k не зависящие от h и такие, что при всех достаточно малых h выполняется оценка*

$$\|\psi_h\|_h\leqslant M_1|h|^k.$$

Введем понятие корректности разностной схемы.

Определение. *Разностная схема (65) называется корректной, если при всех достаточно малых h :*

- ее решение $y_h(x)\in B_h$ существует и единственно при любых правых частях $\psi_h\in B_h$,
- существует постоянная $M_2>0$, не зависящая от h и такая, что при любых правых частях φ_h для решения задачи (65) справедлива оценка

$$\|y_h\|_h\leqslant M_2\|\varphi_h\|_h. \quad (70)$$

Существенным отличием от определения корректности дифференциальной задачи является условие независимости константы M_2 от шагов сетки. Требование 1) в определении корректности означает существование оператора L_h^{-1} , обратного к оператору L_h , а требование 2) — равномерную по h ограниченность оператора L_h^{-1} . Свойство разностной схемы, выраженное неравенством (70), означает непрерывную и равномерную по h зависимость ее решения от правой части. Это свойство называют устойчивостью разностной схемы.

Определение. *Решение разностной задачи (65) сходится к решению исходной задачи (64), (или, более коротко, разностная схема сходится), если*

$$\lim_{h\rightarrow 0}\|z_h\|_h=\lim_{h\rightarrow 0}\|y_h-u_h\|_h=0.$$

Определение. *Говорят, что разностная схема имеет k-й порядок точности, если существует постоянная $M_3>0$, не зависящая от h и постоянная $k>0$ такие, что*

$$\|z_h\|_h\leqslant M_3|h|^k.$$

Покажем, что из аппроксимации и устойчивости разностной схемы следует ее сходимость.

Теорема 10. *(Филлипова). Предположим, что исходная задача (64) поставлена корректно. Пусть разностная схема (65) аппроксимирует исходную задачу (64) и является корректной. Тогда решение разностной задачи (65) сходится к решению исходной задачи (64), причем порядок точности разностной схемы совпадает с порядком аппроксимации.*

Доказательство. Рассмотрим задачу для погрешности (68). Так как уравнение (68) отличается от уравнения (65) только правой частью, из требования устойчивости следует, что для $z_h(x)$ справедлива оценка

$$\|z_h\|_h\leqslant M_2\|\psi_h\|_h.$$

Правая часть этого неравенства стремится к нулю при $|h|\rightarrow 0$, а так как M_2 не зависит от h и, по условию теоремы, разностная схема аппроксимирует исходную задачу. Следовательно $\|z_h\|_h\rightarrow 0$ при $|h|\rightarrow 0$, т.е. схема сходится.

Так как схема имеет *k*-й порядок аппроксимации, то

$$\|\psi_h\|_h\leqslant M_1|h|^k$$

и $\|z_h\|_h\leqslant M_3|h|^k$, $M_3=M_1M_2$, что означает, что схема (65) имеет *k*-й порядок точности. □

Доказанная теорема позволяет разделить исследование сходимости разностной схемы на два этапа: исследование погрешности аппроксимации на решении и получение оценок вида (70), т.е. исследование устойчивости. Как правило, второй этап более трудный, чем первый.

Замечание. *При доказательстве теоремы Филлипова условие согласованности норм (67) не использовалось. Это условие нужно для того, чтобы гарантировать единственность предельной функции. Теорема утверждает, что последовательность y_h сходится к точному решению и $\in B_0$ в том смысле, что $\|y_h-u_h\|_h\rightarrow 0$, при $|h|\rightarrow 0$. Но из теоремы не следует, что не может существовать такая функция $v\in B_0$ (не являющаяся решением задачи (64)), для которой также выполняно условие $\|y_h-v_h\|_h\rightarrow 0$, при $|h|\rightarrow 0$.*

23. Примеры численных методов решения задачи Коши. Погрешность аппроксимации общего 2-этапного метода Рунге-Кутта

$$\begin{cases} \frac{du}{dt}=\mathbf{f}(t,\mathbf{u}(t)),\quad t>0,\\ \mathbf{u}(0)=u_0, \end{cases} \quad (71)$$

где $\mathbf{u}(t)=\left(u_1(t),u_2(t),\ldots,u_m(t)\right)^T$, $\mathbf{f}(t,\mathbf{u}(t))=\left(f_1(t,\mathbf{u}(t)),\ldots,f_m(t,\mathbf{u}(t))\right)^T$.

Пример 1. Пожалуй, наиболее простым методом решения задачи Коши является разностная схема (метод) Эйлера. Несмотря на всю простоту схемы, метод Эйлера часто используется на практике. Метод Эйлера представляет собой разностное уравнение:

$$\begin{cases} \frac{y_{n+1}-y_n}{\tau}=f(t_n,y_n),\quad t_n\in w_\tau\\ y_0=u_0,\quad n\in\mathbb{Z}_+. \end{cases} \quad (72)$$

Эта схема является явной, так как значение численного решения в каждой следующей точке t_{n+1} , $n\in\mathbb{Z}_+$ находится по явной формуле: $y_{n+1}=y_n+\tau f_n$, $n\in\mathbb{Z}_+$. Введем погрешность разностной схемы (72): $z_n=y_n-u_n$, $n\in\mathbb{Z}_+$.

Эта оценка означает, что разностная схема (72) аппроксимирует исходную задачу с первым порядком по τ . В дальнейшем покажем, что рассмотренная разностная схема будет сходиться к решению задачи Коши с первым порядком по τ .

Пример 2. Рассмотрим теперь двухэтапную разностную схему Рунге–Кутта (схему «предиктор–корректор»). В данной разностной схеме вводятся дополнительные точки, так называемые полуцелые слои: $t_{n+\frac{1}{2}}=t_n+0.5\tau$, $n\in\mathbb{Z}_+$.

$$y_{n+1}=y_n+\tau f\left(t_{n+\frac{1}{2}};y_n+0.5\tau f(t_n,y_n)\right). \quad (73)$$

Далее будет показано, что эта двухэтапная разностная схема имеет второй порядок точности по τ .

Оценка погрешности общего двухэтапного метода Рунге–Кутта.

Рассмотрим общий вид двухэтапного метода Рунге–Кутта для уравнения Коши:

$$\begin{cases} \frac{y_{n+1}-y_n}{\tau}=\sigma_1K_1+\sigma_2K_2,\quad n\in\mathbb{Z}_+\\ y_0=u_0,\\ K_1=f(t_n,y_n),\quad K_2=f(t_n+a_2\tau,y_n+b_{21}\tau f(t_n,y_n)), \end{cases} \quad (74)$$

где $\sigma_1,\sigma_2,a_2,b_{21}\in\mathbb{R}$ — некоторые числа, от выбора которых зависит как погрешность аппроксимации, так и точность численного решения.

$\psi_n=-u_n'+(\sigma_1+\sigma_2)f(t_n,u_n)+\tau\left((a_2\sigma_2-0.5)\frac{\partial f_n}{\partial t_n}+(b_{21}\sigma_2-0.5)f_n\frac{\partial f}{\partial x}\right)+O(\tau^2)$.

Потребуем выполнение следующих условий:

- $\sigma_1+\sigma_2=1$ (это условие называется условием аппроксимации).
- $\sigma_2a_2=\sigma_2b_{21}=0.5$.

Тогда погрешность аппроксимации этого метода имеет второй порядок малости по τ : $\psi_n=O(\tau^2)$.

Замечание. *В случае выполнения только первого условия погрешность аппроксимации имеет первый порядок по τ .*

24. Общий *m*-этапный метод Рунге–Кутта. Оценка точности 2-этапного метода.

Рассмотрим задачу Коши для нелинейного обыкновенного дифференциального уравнения первого порядка:

$$\begin{cases} \frac{du}{dt}=f(t,u(t)),\quad t>0\\ u(0)=u_0, \end{cases} \quad (75)$$

где функции $u(t)$ и $f(t,u)$ обладают достаточной гладкостью в соответствующих областях. Считаем, решение $u(t)$ существует и единственно.

$$K_1=f(t_n,y_n),$$

$$K_2=f(t_n+a_2\tau,y_n+b_{21}\tau K_1),$$

$$K_3=f(t_n+a_3\tau,y_n+b_{31}\tau K_1+b_{32}\tau K_2),$$

$$\ldots$$

$$K_m=f(t_n+a_m\tau,y_n+b_{m1}\tau K_1+b_{m2}\tau K_2+\ldots+b_{mm-1}\tau K_{m-1}).$$

При этом разностная схема для исходной задачи (75) имеет вид

$$\begin{cases} \frac{y_{n+1}-y_n}{\tau}=\sigma_1K_1+\sigma_2K_2+\ldots+\sigma_mK_m\\ y_0=u_0,\quad n\in\mathbb{Z}_+, \end{cases} \quad (76)$$

где $\sigma_1,\sigma_2,\ldots,\sigma_m\in\mathbb{R}$.

Будем также считать, что выполнено следующее условие аппроксимации, без которого рассмотрение метода не имеет смысла: $\sum_{i=1}^m\sigma_i=1$.

Замечание. *Заметим, что формулы *m*-этапного метода Рунге–Кутта достаточно громоздки. Это является одной из причин того, что на практике редко используются методы Рунге–Кутта для $m>4$.*

Оценка точности на примере двухэтапного метода Рунге–Кутта. Выпишем еще раз разностную схему, описывающую общий двухэтапный метод Рунге–Кутта:

$$\begin{cases} \frac{y_{n+1}-y_n}{\tau}=(1-\sigma)f(t_n,y_n)+\sigma f(t_n+a\tau,y_n+af(t_n,y_n)),\quad n\in\mathbb{Z}_+\\ y_0=u_0. \end{cases} \quad (77)$$

Введем погрешность разностной схемы (77): $z_n=y_n-u_n$, $n\in\mathbb{Z}_+$. Покажем, что $|z_n|\leqslant M|\psi_n|$, $n\in\mathbb{Z}_+$, где константа M не зависит от шага τ , ψ_n — погрешность аппроксимации на решении исходной задачи Коши:

$$\psi_n=\frac{-u_{n+1}-u_n}{\tau}+(1-\sigma)f(t_n,u_n)+\sigma f(t_n+a\tau,u_n+af(t_n,u_n)).$$

Пусть функция $f(t,u)$ удовлетворяет условию Липшица по второму аргументу с константой $L>0$: $|f(t,v)-f(t,u)|\leqslant L|u-v|$, $(t,u),(t,v)\in G$.

При достаточно малых τ получаем: $|z_{n+1}|=O(\tau^2)$.

25. Многошаговые разностные методы. Понятие устойчивости

Рассмотрим задачу Коши для нелинейного обыкновенного дифференциального уравнения первого порядка:

$$\begin{cases} \frac{du}{dt} = f(t, u(t)), & t > 0, \\ u(0) = u_0, \end{cases} \quad (78)$$

где функции $u(t)$ и $f(t, u)$ обладают достаточной гладкостью. Считаем, что решение $u(t)$ существует и единственно.

Определение. *Линейным m -шаговым разностным методом решения задачи (78) называется разностная схема вида*

$$\sum_{k=0}^m \frac{a_k}{\tau} y_{n-k} = \sum_{k=0}^m b_k f_{n-k}, \quad (79)$$

где $m \in \mathbb{N}$, $\tau > 0$ – шаг сетки ω_τ , $a_k, b_k \in \mathbb{R}$, $k = \overline{0, m}$, причем $a_0 \neq 0, b_m \neq 0$.

Замечание. Уравнение (79) следует рассматривать как рекуррентное соотношение, выражающее новое значение $y_n = y(t_n)$ через найденные ранее значения $y_{n-1}, y_{n-2}, \dots, y_{n-m}$. Уравнение (79) определено для $n = m, m+1, \dots$ и требует для начала расчета задания m начальных значений $y_0, y_1, \dots, y_{m-1}$. Значение $y_0 = u(0)$ определяется исходной задачей (78), а величины $y_1, \dots, y_{m-1}$ можно вычислить с помощью других методов, например, с помощью рассмотренного выше метода Рунге–Кутты. В дальнейшем будем предполагать, что величины $y_0, y_1, \dots, y_{m-1}$ уже заданы.

Если в разностной схеме (79) $b_0 = 0$, то рассматриваемый метод называется явным, и искомое значение y_n выражается явным образом через предыдущие: $\frac{a_0}{\tau} y_n = \sum_{k=1}^m b_k f_{n-k} - \sum_{k=1}^m \frac{a_k}{\tau} y_{n-k}$.

Если $b_0 \neq 0$, то метод называется неявным, и для нахождения y_n приходится решать нелинейное уравнение $\frac{a_0}{\tau} y_n - b_0 f(t_n, y_n) = F(y_{n-1}, \dots, y_{n-m})$, где $F(y_{n-1}, \dots, y_{n-m}) = \sum_{k=1}^m (b_k f_{n-k} - \frac{a_k}{\tau} y_{n-k})$. Обычно это уравнение решают итерационным методом Ньютона, выбирая начальное приближение равным y_{n-1} . Заметим, что коэффициенты уравнения (79) определены с точностью до множителя. Для определенности будем считать, что выполнено условие $\sum_{k=0}^m b_k = 1$.

Определение. *Погрешностью аппроксимации разностной схемы (79) на решении исходной задачи (78) называется сеточная функция*

$$\psi_n = - \sum_{k=0}^m \frac{a_k}{\tau} u_{n-k} + \sum_{k=0}^m b_k f(t_{n-k}, u_{n-k}), \quad (80)$$

заданная на сетке ω_τ , где $u_n = u(t_n)$ – решение исходной задачи (78).

Для сходимости окончательно получаем следующую систему уравнений:

$$\begin{cases} \sum_{k=1}^m ka_k = -1, \\ \sum_{k=0}^m k^{l-1} (ka_k + lb_k) = 0, & l = \overline{2, p}, \end{cases} \quad (81)$$

в которой коэффициенты a_0, b_0 вычисляются по формулам $a_0 = -\sum_{k=1}^m a_k$, $b_0 = 1 - \sum_{k=1}^m b_k$. Таким образом, мы уменьшили число уравнений в системе до p и число неизвестных до $2m$. Чтобы система не была переопределенной (в таких системах число уравнений больше числа неизвестных) необходимо выполнение условия $p \leq 2m$.

Таким образом наибольший возможный порядок аппроксимации неявных m -шаговых разностных методов равен $2m$, явных – $(2m-1)$, так как в явных методах $b_0 = 0$, и число неизвестных в системе (81) меньше на единицу по сравнению с системой, записанной для неявного метода.

Замечание 20. Если убрать последние p уравнений системы (81), $n = 1, (p-1)$, то получим условия, обеспечивающие порядок погрешность аппроксимации $O(\tau^{p-n})$.

Замечание 21. В практике вычислений наибольшее распространение получили методы Адамса, которые представляют собой частный случай многошаговых методов (79), когда производная $u'(t)$ в исходном уравнении аппроксимируется по двум крайним точкам t_{n-1} и t_n , то есть $a_0 = 1$, $a_1 = -1$, $a_k = 0$, $k = 2, \overline{m}$:

$$\frac{y_n - y_{n-1}}{\tau} = \sum_{k=0}^m b_k f_{n-k}.$$

Замечание 22. Разностные схемы вида (79), обладающие наименьшими порядками аппроксимации на решении исходного уравнения, неустойчивы и не могут быть использованы на практике. Максимальный порядок аппроксимации устойчивого неявного m -шагового метода не превосходит $(m+1)$, если m нечетно, и не превосходит $(m+2)$, если m четно. Порядок аппроксимации устойчивых явных схем не превосходит m . Подробнее понятие устойчивости m -шагового разностного метода мы рассмотрим в следующем параграфе.

Достоинства и недостатки многошаговых разностных методов по сравнению с методом Рунге–Кутты.

Достоинства:

- Формулы многошаговых методов значительно проще.
- Многошаговые методы позволяют достигать большей точности.

Недостатки:

- В многошаговых методах необходимо хранить в памяти большее число элементов – значения нескольких предыдущих шагов вместо одного.
- Многошаговые методы требуют наличия «разгонного этапа», то есть значений нескольких первых шагов, которые нельзя вычислить по многошаговым формулам. Как мы уже упоминали, эти значения обычно вычисляют с помощью метода Рунге–Кутты.

Понятие устойчивости разностного метода

Численный метод называется устойчивым, если погрешности, допущенные на каком-то этапе вычислений, не оказывают существенного влияния на результат.

$y_{n+1} = qy_n$, $\tilde{y}_n = y_n + \delta_n$.

$$\tilde{y}_{n+1} = q\tilde{y}_n = qy_n + q\delta_n = y_{n+1} + q\delta_n,$$

$$\begin{cases} \frac{du(t)}{dt} + \lambda u(t) = 0, & \lambda > 0, \ t > 0, \\ u(0) = u_0. \end{cases} \quad (82)$$

Пример 3. Рассмотрим, например, явную разностную схему Эйлера для решения задачи (82):

$$\frac{y_{n+1} - y_n}{\tau} + \lambda y_n = 0, \lambda > 0,$$

В этом случае схема называется условно устойчивой, а само неравенство $\tau \leq \frac{2}{\lambda}$ называется условием устойчивости.

Пример 4. Приведем пример абсолютно устойчивой разностной схемы. Для уравнения

$$u'(t) = f(t, u(t)), \quad (83)$$

$y_{n+1} = qy_n$, $q = (1 + \tau\lambda)^{-1}$, причем $|q| < 1$ при любых $\tau > 0$.

Исследуем на устойчивость двухшаговый разностный метод, построенный в примере 3

$$\frac{y_{n+1} - y_n}{\tau} = \frac{3}{2}f_n - \frac{1}{2}f_{n-1}. \quad (84)$$

$\frac{y_{n+1} - y_n}{\tau} + \lambda(\frac{3}{2}f_n - \frac{1}{2}f_{n-1}) = 0$, $n = 1, 2, \dots$, представляет собой разностное уравнение второго порядка с постоянными коэффициентами

$$y_{n+1} + ay_n + by_{n-1} = 0, \quad (85)$$

Лемма 1. *Оба корня уравнения (85) с действительными коэффициентами a, b лежат внутри или на границе единичного круга $|q| \leq 1$ тогда и только тогда, когда выполнены условия*

$$1 + a + b \geq 0, \ 1 - a + b \geq 0, \ b \leq 1. \quad (86)$$

Общий m -шаговый линейный разностный метод

$$\sum_{k=0}^m \frac{a_k}{\tau} y_{n-k} = \sum_{k=0}^m b_k f_{n-k}, \quad (87)$$

где $\tau > 0$, $y_0, y_1, \dots, y_{m-1}$ – заданы. Будем считать, что коэффициенты a_k, b_k , $k = \overline{1, n}$ не зависят от τ .

Пример. В применении к уравнению (82) метод (87) принимает вид:

$$\sum_{k=0}^m (a_k + \tau\lambda b_k) y_{n-k} = 0. \quad (88)$$

Решение этого разностного уравнения с постоянными коэффициентами будем искать в виде $y_j = q^j$, $j \in \mathbb{Z}_+$.

$$F_m(q, \tau) = \sum_{k=0}^m (a_k + \tau\lambda b_k) q^{m-k} = 0. \quad (89)$$

Определение. *Уравнение вида (89) называется характеристическим уравнением разностной схемы (88).*

$F_m(q, 0) = 0$,

$$\sum_{k=0}^m a_k q^{m-k} = 0, \quad (90)$$

Определение. *Говорят, что схема (87) удовлетворяет условию (α) , если все корни характеристического уравнения (90) лежат внутри или на границе единичного круга комплексной плоскости, причем на границе единичного круга нет кратных корней.*

Таким образом, выполнение условия (α) соответствует устойчивости разностного метода для уравнения $u'(t) = 0$. Однако часто схему и для общего уравнения (83) называют устойчивой, если она удовлетворяет условию (α) . Такая терминологическая неточность оправдана тем, что из условия (α) следует сходимость решения разностной задачи (87) к решению исходной дифференциальной задачи (83).

Теорема. *Пусть разностная схема удовлетворяет условию (α) и $|f'_u| \leq L$ на отрезке $0 \leq t \leq T$. Тогда при $0 \leq t_n = n\tau \leq T$ и всех достаточно малых τ выполняется оценка*

$$|y_n - u(t_n)| \leq M \left(\sum_{j=m}^n \tau |\psi_j| + \max_{0 \leq i \leq n-1} |y_i - u(t_i)| \right),$$

где $|y_i - u(t_i)|$ – погрешности в задании начальных данных, $i = 0, (m-1), M$ – константа, зависящая от L, T и не зависящая от τ , ψ_j – погрешность аппроксимации на решении исходного уравнения (83):

$$\psi_j = - \sum_{k=0}^m \frac{a_k}{\tau} u(t_{n-k}) + \sum_{k=0}^m b_k f_{n-k}.$$

(без доказательства)

Замечание 23. Методы Адамса $\frac{y_n - y_{n-1}}{\tau} = \sum_{k=0}^m b_k f_{n-k}$ всегда удовлетворяют условию (α) , так как для них $a_0 = -a_1 = 1$, то есть $q = q_1 = 1$, что следует из уравнения $q^n - q^{n-1} = 0$.

Замечание 24. При указанном подходе, в отличие от рассмотренных примеров, не различаются абсолютно устойчивые и условно устойчивые разностные схемы, так как параметр τ заранее считается достаточно малым.

Замечание 25. Мы уже упоминали в ?? данной главы, что наивысший достижимый порядок аппроксимации неявных m -шаговых методов равен $2m$, а явных – $(2m-1)$. Однако оказывается, что методы наивысшего порядка неустойчивы в том смысле, что они не удовлетворяют условию (α) . А именно, если m нечетно, то никакой устойчивый метод не превосходит порядка $p = m+1$. Если m четно, то никакой устойчивый метод не превосходит порядка $p = m+2$ (p – порядок аппроксимации). Для явных схем наивысший порядок аппроксимации устойчивых методов $p = m$.

Пример. Нетрудно привести пример схем, не удовлетворяющих условию (α) . Так, явная двухшаговая схема

$$\frac{y_n + 4y_{n-1} - 5y_{n-2}}{6\tau} = \frac{2f_{n-1} + f_{n-2}}{3}$$

имеет третий порядок погрешности аппроксимации $\psi = O(\tau^3)$

26. Жесткие системы ОДУ.

Определение. *Система линейных обыкновенных дифференциальных уравнений вида*

$$\begin{cases} \frac{du(t)}{dt} = Au(t), & t > 0 \\ u(0) = u_0, \end{cases}$$

где $\mathbf{u}(t) = (u_1(t), u_2(t), \dots, u_m(t))^T$, и A ($m \times m$) – заданная матрица постоянных, вообще говоря, комплексных коэффициентов, называется жесткой, если:

- Действительные части всех собственных значений λ_k , $k = \overline{1, m}$ матрицы A отрицательные.
- Выполняется неравенство $\frac{\max_{1 \leq k \leq m} |Re\lambda_k|}{\min_{1 \leq k \leq m} |Re\lambda_k|} \gg 1$.

Понятие жесткости можно обобщить и на случай нелинейных систем. Рассмотрим систему нелинейных обыкновенных дифференциальных уравнений

$$\begin{cases} \frac{du(t)}{dt} = \mathbf{f}(t, \mathbf{u}(t)), & t > 0, \\ \mathbf{u}(0) = u_0, \end{cases} \quad (91)$$

где $\mathbf{u}(t) = (u_1(t), u_2(t), \dots, u_m(t))^T$, $\mathbf{f}(t, \mathbf{u}(t)) = (f_1(t, \mathbf{u}(t)), f_2(t, \mathbf{u}(t)), \dots, f_m(t, \mathbf{u}(t)))^T$.

Пусть $\lambda_k(t)$, $k = \overline{1, m}$ – собственные значения матрицы $J(t) = \frac{\partial \mathbf{f}(t, \mathbf{v}(t))}{\partial \mathbf{u}}$.

$$\text{Введем число жесткости } S(t) = \frac{\max_{1 \leq k \leq m} |Re\lambda_k(t)|}{\min_{1 \leq k \leq m} |Re\lambda_k(t)|}.$$

Определение. *Система (91) называется жесткой на решении $\mathbf{v}(t)$ и на данном интервале $0 < t < T$ если*

- $Re\lambda_k(t) < 0$, $k = \overline{1, m}$.
- Число жесткости $S(t)$ велико на рассматриваемом интервале $0 < t < T$: $\frac{\max_{1 \leq k \leq m} |Re\lambda_k(t)|}{\min_{1 \leq k \leq m} |Re\lambda_k(t)|} \gg 1$.

Заметим, что первое требование означает асимптотическую устойчивость по Ляпунову решения $\mathbf{v}(t)$.

Дальнейшие определения устойчивости

$$\frac{d\mathbf{u}(t)}{dt} = \lambda \mathbf{u}(t), \quad (92)$$

Определение. *Разностный метод называется $A(\alpha)$ -устойчивым, если область его устойчивости содержит угол левой полуплоскости: $|\arg(-\mu)| < \alpha$, $\mu = \tau\lambda, \alpha > 0$. В частности, $A(\frac{\pi}{2})$ -устойчивость совпадает с A -устойчивостью.*

Известно, что ни для какого α не существует явного $A(\alpha)$ -устойчивого линейного многошагового метода. Построены $A(\alpha)$ -устойчивые неявные методы третьего и четвертого порядка аппроксимации. К ним относятся чисто неявные многошаговые разностные схемы, у которых правая часть $\mathbf{f}(t, \mathbf{u})$ вычисляется только при новом значении $t = t_{n+m}$, а производная $\mathbf{u}'(t)$ аппроксимируется по нескольким предыдущим точкам и точке $t = t_{n+m}$.

hellow 4