

ЭЛЛИПТИЧЕСКИЕ УРАВНЕНИЯ

Глава XII посвящена методам решения краевых задач для эллиптических уравнений. В § 1 решение таких задач сводится к решению эволюционных задач для параболических уравнений до выхода на стационарный режим; последнее выполняется при помощи многомерных разностных схем, изложенных в гл. XI, § 2. Обсужден выбор оптимального шага по времени (или набора переменных шагов) в таких расчетах.

В § 2 рассмотрены вариационные методы решения эллиптических уравнений и вариационные способы составления стационарных (не эволюционных) разностных схем. В последнем случае указаны прямые и итерационные методы вычисления разностного решения.

§ 1. Счет на установление

1. Стационарные решения эволюционных задач. К эллиптическим уравнениям приводит ряд физических задач: определение прогиба нагруженной мембраны, давления газа в неоднородном силовом поле, стационарного (не зависящего от времени) распределения тепла в теле и т. д. Все эти задачи имеют общее свойство: предполагается, что внешние воздействия не зависят от времени, а начальные условия были заданы достаточно давно, так что физическая система успела выйти на стационарное решение $u(\mathbf{r})$, не зависящее от времени.

Примером полной математической постановки является задача с краевыми условиями первого рода, называемая задачей Дирихле; требуется найти непрерывное решение задачи

$$\Delta u(\mathbf{r}) = -f(\mathbf{r}), \quad \mathbf{r} \in G, \quad u_{\Gamma}(\mathbf{r}) = \mu(\mathbf{r}), \quad (1)$$

где $G(\mathbf{r})$ есть многомерная замкнутая область с границей Γ . В отличие от эволюционных задач, разобранных в предыдущих главах, постановка (1) не содержит начальных условий. Обобщением задачи (1) является следующая задача:

$$\operatorname{div}[k(\mathbf{r}) \operatorname{grad} u(\mathbf{r})] = -f(\mathbf{r}), \quad \mathbf{r} \in G, \quad u_{\Gamma}(\mathbf{r}) = \mu(\mathbf{r}), \quad k(\mathbf{r}) > 0. \quad (2)$$

Задачи с другими краевыми условиями мы не будем рассматривать.

Задачу (2) будем называть *стационарной*. Наряду с ней рассмотрим *эволюционную* задачу для параболического уравнения с теми же граничными условиями и произвольно выбранными начальными данными:

$$\frac{\partial v(\mathbf{r}, t)}{\partial t} = \operatorname{div} [k(\mathbf{r}) \operatorname{grad} v(\mathbf{r}, t)] + f(\mathbf{r}), \quad \mathbf{r} \in G, \quad 0 \leq t < +\infty, \quad (3)$$

$$v_{\Gamma}(\mathbf{r}, t) = \mu(\mathbf{r}), \quad v(\mathbf{r}, 0) = v_0(\mathbf{r}).$$

Исследуем, насколько решение $v(\mathbf{r}, t)$ эволюционной задачи отличается от решения $u(\mathbf{r})$ стационарной задачи. Вычитая (2) из (3) и учитывая, что $\partial u(\mathbf{r})/\partial t = 0$, найдем, что разность $w(\mathbf{r}, t) = v(\mathbf{r}, t) - u(\mathbf{r})$ удовлетворяет однородному параболическому уравнению с однородными краевыми условиями:

$$\frac{\partial w(\mathbf{r}, t)}{\partial t} = \operatorname{div} [k(\mathbf{r}) \operatorname{grad} w], \quad \mathbf{r} \in G, \quad 0 < t < +\infty, \quad (4)$$

$$w_{\Gamma}(\mathbf{r}, t) = 0, \quad w(\mathbf{r}, 0) = w_0(\mathbf{r}) \equiv v_0(\mathbf{r}) - u(\mathbf{r}).$$

Поскольку начальные данные в (3) были выбраны произвольно, то без ограничения общности можно считать, что начальные данные задачи (4) также выбраны произвольно.

В курсах математической физики показано (см., например, [40]), что при помощи метода разделения переменных решение задачи (4) можно представить в следующем виде:

$$w(\mathbf{r}, t) = \sum_{q=1}^{\infty} c_q e^{-\lambda_q t} w_q(\mathbf{r}). \quad (5)$$

Здесь $w_q(\mathbf{r})$ и λ_q — собственные функции и собственные значения многомерной задачи Штурма — Лиувилля:

$$\operatorname{div} [k(\mathbf{r}) \operatorname{grad} w_q] + \lambda_q w_q(\mathbf{r}) = 0, \quad \mathbf{r} \in G, \quad w_q(\mathbf{r})_{\Gamma} = 0, \quad (6)$$

а

$$c_q = \int_G w(\mathbf{r}, 0) w_q(\mathbf{r}) d\mathbf{r}$$

являются коэффициентами Фурье начальных данных (4) по системе функций $w_q(\mathbf{r})$. Собственные значения задачи (6) положительны и образуют неубывающую последовательность

$$0 < \lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \dots, \quad (7)$$

а собственные функции $w_q(\mathbf{r})$ образуют полную ортонормированную систему в $G(\mathbf{r})$.

Из (5) и (7) нетрудно получить неравенство

$$\| \omega(\mathbf{r}, t) \|_{L_2} = \left(\sum_{q=1}^{\infty} c_q^2 e^{-2\lambda_q t} \right)^{1/2} \leq e^{-\lambda_1 t} \left(\sum_{q=1}^{\infty} c_q^2 \right)^{1/2} = e^{-\lambda_1 t} \| \omega_0(\mathbf{r}) \|_{L_2}. \quad (8)$$

Оно означает, что разность $\omega(\mathbf{r}, t) = v(\mathbf{r}, t) - u(\mathbf{r})$ при $t \rightarrow \infty$ экспоненциально стремится к нулю по норме $\| \cdot \|_{L_2}$, так что решение $v(\mathbf{r}, t)$ эволюционной задачи (3) среднеквадратично сходится к решению $u(\mathbf{r})$ стационарной задачи (2) при $t \rightarrow \infty$.

Замечание 1. Пусть граничные и начальные условия таковы, что решения задач (2) и (3) имеют в $G(\mathbf{r})$ непрерывные производные, ограниченные равномерно по t . Тогда сходимость $v(\mathbf{r}, t)$ к $u(\mathbf{r})$ будет равномерной.

Таким образом, вместо задачи (2) для эллиптического уравнения можно взять эволюционную задачу (3) для параболического уравнения с тем же пространственным оператором, произвольно выбрать начальные данные и вычислить решение $v(\mathbf{r}, t)$ при достаточно большом t . Стационарный (не зависящий от времени) предел $u(\mathbf{r})$, к которому стремится $v(\mathbf{r}, t)$ при $t \rightarrow \infty$, и будет решением стационарной задачи (2).

Этот способ называется *счетом на установление*. Он позволяет осуществить численное решение эллиптических задач хорошо разработанными методами решения параболических задач, например, продольно-поперечной схемой для двумерных задач и локально-одномерными схемами в случае большего числа измерений.

Установление стационарного решения происходит довольно быстро благодаря экспоненциальному характеру затухания начальных данных. Из (8) видно, что если нужна точность $\sim \varepsilon$, то надо вести вычисления до момента

$$T \approx \frac{1}{\lambda_1} \ln \frac{1}{\varepsilon}, \quad (9)$$

где λ_1 есть наименьшее собственное значение соответствующей задачи Штурма — Лиувилля (6).

Замечание 2. На стационарное решение выходят не только решения параболических задач. То же происходит при других диссипативных процессах со стационарными граничными условиями, например при колебаниях с вязким трением, описываемых уравнением

$$v_{tt} + \beta v_t = \operatorname{div}(k \operatorname{grad} v) + f, \quad v_{\Gamma} = \mu(\mathbf{r}), \quad \beta > 0. \quad (10)$$

Можно формулировать эволюционную задачу для этого уравнения; однако это менее удобно.

Замечание 3. Можно составить разностную схему, непосредственно аппроксимирующую исходную задачу (2). Но в § 2

мы увидим, что вычислять разностное решение при этом обычно приходится итерационными методами. Оказывается, что соответствующие итерационные алгоритмы можно интерпретировать как некоторые разностные схемы для эволюционной задачи (3).

2. Оптимальный шаг. Для расчета эволюционной p -мерной задачи (3) до момента T используют экономичные разностные схемы. При этом шаги τ и h_α ($1 \leq \alpha \leq p$) выбирают достаточно малыми, чтобы обеспечить требуемую близость разностного решения y к точному решению $v(\mathbf{x}, t)$ эволюционной задачи.

Однако если шаг τ выбран слишком малым, то расчет до момента T потребует большого числа шагов, что неоправданно увеличит объем вычислений. Очевидно, должен существовать оптимальный шаг τ_0 ; рассмотрим, как его найти.

Для простоты ограничимся двумерной задачей Дирихле в прямоугольнике:

$$\begin{aligned} u_{x_1 x_1} + u_{x_2 x_2} &= -f(\mathbf{x}), \quad 0 < x_1 < a, \quad 0 < x_2 < b, \\ u_\Gamma(\mathbf{x}) &= \mu(\mathbf{x}), \quad \mathbf{x} = \{x_1, x_2\}. \end{aligned} \quad (11)$$

Ей соответствует эволюционная задача для уравнения

$$v_t = v_{x_1 x_1} + v_{x_2 x_2} + f(\mathbf{x}), \quad (12)$$

которую будем решать на равномерной сетке $\{x_{1n} = nh_1, x_{2m} = mh_2, 0 \leq n \leq N, 0 \leq m \leq M\}$ с шагами $h_1 = a/N, h_2 = b/M$.

Продольно-поперечная схема. Для исследования этой схемы возьмем ее запись (11.63) в двуслойной форме:

$$\frac{1}{\tau}(\hat{y} - y) = \frac{1}{2}(\Lambda_1 + \Lambda_2)(\hat{y} + y) - \frac{\tau}{4}\Lambda_1\Lambda_2(\hat{y} - y) + \bar{f},$$

и преобразуем ее к канонической форме:

$$B \frac{\hat{y} - y}{\tau} + Ay = \varphi, \quad \varphi = \bar{f}, \quad (13a)$$

где

$$A = -(\Lambda_1 + \Lambda_2), \quad B = \left(E - \frac{\tau}{2}\Lambda_1\right)\left(E - \frac{\tau}{2}\Lambda_2\right). \quad (13b)$$

Поскольку в уравнении (12) коэффициент теплопроводности $k = 1$, а сетка равномерна, то

$$\begin{aligned} \Lambda_1 y_{nm} &= \frac{1}{h_1^2} (y_{n+1, m} - 2y_{nm} + y_{n-1, m}), \\ \Lambda_2 y_{nm} &= \frac{1}{h_2^2} (y_{n, m+1} - 2y_{nm} + y_{n, m-1}). \end{aligned} \quad (13b)$$

Если численный расчет доведен до выхода на стационарное решение, то $\hat{y} \approx y$. Тогда схема (13) в пределе переходит в не-

эволюционную (не содержащую времени) разностную схему

$$Ay = \varphi, \quad A = -(\Lambda_1 + \Lambda_2), \quad \varphi = f(x), \quad (14)$$

которая, как нетрудно заметить, аппроксимирует стационарную задачу (11). Очевидно, в этом случае оптимальным будет тот шаг τ_0 , при котором разностное решение выйдет на стационарное за наименьшее число шагов. Для этого надо, чтобы начальные данные за один шаг затухали возможно сильнее.

Затухание начальных данных можно исследовать методом разделения переменных, взятым в строгой форме (поскольку нас интересуют точные значения границ спектра оператора). Собственные функции разностного оператора $-(\Lambda_1 + \Lambda_2)$ в прямоугольнике на равномерной сетке равны, как нетрудно проверить,

$$\begin{aligned} \omega_{qr}(x) &= \sin \frac{\pi q x_1}{a} \sin \frac{\pi r x_2}{b}, \\ 1 \leq q \leq N-1, \\ 1 \leq r \leq M-1. \end{aligned} \quad (15)$$

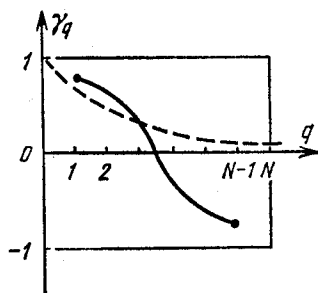


Рис. 86.

Подставляя их в схему (13) и полагая $\hat{w}_{qr} = \rho_{qr} \omega_{qr}$, определим множители роста гармоник:

$$\begin{aligned} \rho_{qr} &= \frac{(1 - \alpha_q)(1 - \beta_r)}{(1 + \alpha_q)(1 + \beta_r)}, \\ \alpha_q &= \frac{2\tau}{h_1^2} \sin^2 \frac{\pi q h_1}{2a}, \quad \beta_r = \frac{2\tau}{h_2^2} \sin^2 \frac{\pi r h_2}{2b}. \end{aligned} \quad (16)$$

Очевидно, все $|\rho_{qr}| < 1$, т. е. все гармоники затухают; это означает, что схема (13) обладает аппроксимационной вязкостью.

Какие гармоники затухают наиболее медленно и, тем самым, сильнее всего препятствуют выходу на стационарный режим? Нетрудно заметить, что входящий в ρ_{qr} сомножитель $\gamma_q = (1 - \alpha_q)/(1 + \alpha_q)$ заключен в пределах $(-1, +1)$ и монотонно убывает при увеличении номера q (рис. 86, жирная линия). Наибольшим по модулю может быть множитель либо с $q=1$, либо с $q=N-1$. Считая N достаточно большим, можно положить

$$\sin \frac{\pi h_1}{2a} = \sin \frac{\pi}{2N} \approx \frac{\pi}{2N}, \quad \sin \frac{\pi(N-1)h_1}{2a} = \sin \frac{\pi(N-1)}{2N} \approx 1$$

и представить экстремальные множители (при $\tau \sim h$) в виде

$$\gamma_1 \approx 1 - \frac{\pi^2 \tau}{a^2}, \quad \gamma_{N-1} \approx -\left(1 - \frac{a^2}{\tau N^2}\right). \quad (17)$$

Аналогично, второй сомножитель $(1 - \beta_r)/(1 + \beta_r)$ максимален по модулю либо при $r=1$, либо при $r=M-1$. Поэтому $|\rho_{qr}|$ максимален либо при $q=r=1$, либо при $q=N-1, r=M-1$, причем

$$\rho_{11} \approx 1 - \pi^2 \tau \left(\frac{1}{a^2} + \frac{1}{b^2} \right), \quad \rho_{N-1, M-1} \approx 1 - \frac{a^2}{\tau N^2} - \frac{b^2}{\tau M^2}. \quad (18)$$

Чем больше шаг τ , тем меньше ρ_{11} и больше $\rho_{N-1, M-1}$, причем оба они близки к 1*); это значит, что первая и последняя гармоники затухают медленно, причем при малом шаге τ быстрее затухает последняя, а при большом — первая гармоника. Выберем шаг τ_0 так, чтобы $\rho_{11}(\tau_0) = \rho_{N-1, M-1}(\tau_0)$; из (18) видно, что

$$\tau_0 \approx \frac{1}{\pi} \left(\frac{a^2}{N^2} + \frac{b^2}{M^2} \right)^{1/2} \left(\frac{1}{a^2} + \frac{1}{b^2} \right)^{-1/2}. \quad (19)$$

Если изменить шаг по сравнению с τ_0 , то либо первая, либо последняя гармоника будет затухать медленнее, чем при $\tau = \tau_0$. Следовательно, τ_0 есть оптимальный шаг.

Число шагов $K(\varepsilon)$, нужное для достижения заданной точности ε , определяется условием $(\max |\rho_{qr}|)^K = \varepsilon$. При оптимальном шаге наибольшие множители роста равны

$$\rho_{11}(\tau_0) = \rho_{N-1, M-1}(\tau_0) \approx \exp \left[-\pi \left(\frac{a^2}{N^2} + \frac{b^2}{M^2} \right)^{1/2} \left(\frac{1}{a^2} + \frac{1}{b^2} \right)^{1/2} \right]. \quad (20)$$

Поэтому минимально необходимое число шагов есть

$$K(\tau_0) = \frac{1}{\pi} \left(\frac{a^2}{N^2} + \frac{b^2}{M^2} \right)^{-1/2} \left(\frac{1}{a^2} + \frac{1}{b^2} \right)^{-1/2} \ln \frac{1}{\varepsilon}. \quad (21)$$

Сравнивая время счета на установление (9) и величину оптимального шага (19), нетрудно убедиться, что $K(\tau_0) = T/\tau_0$.

Отметим, что при $a=b, M=N$ выполняется $\tau_0 \sim a^2/N$ и $K(\tau_0) \sim N$.

В дифференциальном уравнении (12) установление происходит за достаточно большой промежуток времени. Почему же не взять для разностной схемы очень большой шаг по времени, если устойчивость это позволяет? Кажется бы, тогда мы быстрее добьемся установления. Но это не так. Спектр дифференциального оператора таков, что гармоники затухают тем быстрее, чем больше их номер, причем $|\rho_q| \rightarrow 0$ при $q \rightarrow \infty$; соответствующая кривая показана пунктиром на рис. 86. А затухание собственных функций разностного оператора имеет, вообще говоря, другой качественный характер, как видно из того же рисунка.

*) При нечетном числе измерений p для последней гармоники $\rho \rightarrow -1$ и вместо него во всех выкладках надо использовать $|\rho| = -\rho$.

Локально-одномерная схема (11.69) с полусуммой по времени в двумерном случае может быть записана в виде

$$\left(E - \frac{\tau}{2} \Lambda_1\right) \bar{y} = \left(E + \frac{\tau}{2} \Lambda_1\right) y + \tau \varphi_1, \quad (22a)$$

$$\left(E - \frac{\tau}{2} \Lambda_2\right) \hat{y} = \left(E + \frac{\tau}{2} \Lambda_2\right) \bar{y} + \tau \varphi_2, \quad (22б)$$

где операторы Λ_α имеют вид (13в) и коммутируют друг с другом. Умножая уравнение (22a) слева на $(E + \frac{1}{2}\tau\Lambda_2)$, а уравнение (22б) — на $(E - \frac{1}{2}\tau\Lambda_1)$, исключим \bar{y} и запишем (22) в виде двуслойной схемы:

$$\begin{aligned} \left(E - \frac{\tau}{2} \Lambda_1\right) \left(E - \frac{\tau}{2} \Lambda_2\right) \hat{y} = \\ = \left(E + \frac{\tau}{2} \Lambda_1\right) \left(E + \frac{\tau}{2} \Lambda_2\right) y + \tau (\varphi_1 + \varphi_2) + \frac{\tau^2}{2} (\Lambda_2 \varphi_1 - \Lambda_1 \varphi_2). \end{aligned}$$

Преобразуем ее к канонической форме:

$$\begin{aligned} \left(E - \frac{\tau}{2} \Lambda_1\right) \left(E - \frac{\tau}{2} \Lambda_2\right) \frac{\hat{y} - y}{\tau} - (\Lambda_1 + \Lambda_2) y = \\ = \varphi_1 + \varphi_2 + \frac{\tau}{2} (\Lambda_2 \varphi_1 - \Lambda_1 \varphi_2). \quad (23) \end{aligned}$$

Видно, что левая часть (23) совпадает с левой частью продольно-поперечной схемы (13). Поэтому шаг τ_0 , обеспечивающий наиболее быстрое затухание начальных данных, для схемы (23) определяется также формулой (19).

Нетрудно понять, как обобщить выражения оптимального шага (19) и минимального числа шагов (21) на случай произвольного числа измерений для локально-одномерной схемы с полусуммой. Запишем эти выражения в простейшем случае, когда задача Дирихле поставлена в p -мерном кубе со стороной a и по каждой стороне взято N узлов сетки:

$$\tau_0 = \frac{a^2}{\pi N}, \quad K(\tau_0) = \frac{N}{\pi p} \ln \frac{1}{\varepsilon}. \quad (24)$$

Однако если положить $\varphi_1 + \varphi_2 = f(x)$, то правая часть (23) будет отличаться от $f(x)$ на величину $O(\tau)$. Поэтому установившееся разностное решение будет отличаться от $u(x)$ на $O(\tau + h^2)$, и, тем самым, общая точность расчета будет хуже, чем по продольно-поперечной схеме.

Замечание. Для улучшения точности приравняем $f(x)$ правой части (23). Для этого достаточно произвольно выбрать $\varphi_1(x)$, а $\varphi_2(x)$ определить из уравнения

$$\varphi_2 - \frac{\tau}{2} \Lambda_1 \varphi_2 = f - \varphi_1 - \frac{\tau}{2} \Lambda_2 \varphi_1. \quad (25)$$

Это линейное уравнение с трехдиагональной матрицей; оно легко решается одномерной прогонкой по направлению x_1 .

Произвольная область. Выбрать оптимальный шаг удастся только в простейших задачах, когда точно известны границы спектра разностного оператора. В областях сложной формы мы можем, подставляя в формулу (19) характерные размеры области и число узлов сетки, определить лишь порядок величины τ_0 . Поэтому надо представлять, как зависит число шагов $K(\tau)$, требуемое для установления с заданной точностью, от величины шага τ .

При $\tau \leq \tau_0$ затухание начальных данных определяется множителем ρ_{11} , так что число шагов $K_\varepsilon(\tau)$ находится из условия $\rho_{11}^K = \varepsilon$. Из формулы (18) с учетом малости τ следует, что $\rho_{11}(\tau) \approx \exp(-\text{const} \cdot \tau)$; тем самым,

$$\rho_{11}(\tau) \approx [\rho_{11}(\tau_0)]^{\tau/\tau_0}.$$

Отсюда нетрудно получить, что

$$K_\varepsilon(\tau) = \frac{\tau_0}{\tau} K(\tau_0) \quad \text{при} \quad \tau \leq \tau_0. \quad (26a)$$

Аналогично находим

$$K_\varepsilon(\tau) = \frac{\tau}{\tau_0} K(\tau_0) \quad \text{при} \quad \tau_0 \leq \tau. \quad (26b)$$

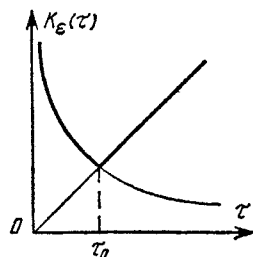


Рис. 87.

Кривая $K_\varepsilon(\tau)$ изображена на рис. 87 жирной линией. Ее минимум соответствует оптимальному шагу. Видно, что в нижней точке кривая имеет разрыв производной. Значит, умеренное отклонение величины шага τ от оптимума может заметно увеличить требуемое число шагов (во столько раз, во сколько τ отличается от τ_0).

Критерии установления. Из сказанного выше следует, что для задач в достаточно общей постановке (2), (3) заранее неизвестно, какое число шагов надо сделать до установления. Поэтому на практике вычисления прекращают при выполнении какого-нибудь правдоподобного, хотя и нестрогого критерия.

Нередко пользуются простейшим критерием

$$\|\hat{y} - y\| \leq \varepsilon; \quad (27a)$$

однако он недостаточно надежен, поскольку разностное решение устанавливается медленно. Если учесть, что установление происходит почти по геометрической прогрессии, то нетрудно получить более надежный критерий:

$$\|\hat{y} - y\| \leq \varepsilon(1 - \nu), \quad \text{где} \quad \nu = \frac{\|\hat{y} - y\|}{\|y - \hat{y}\|}. \quad (27b)$$

Для схемы типа (13) расчет иногда оканчивают по условию малости невязки:

$$\|Ay - f\| \leq \varepsilon. \quad (27в)$$

Комплексная организация расчета, описанная в гл. VIII, § 2, п. 5, очень полезна даже в одномерных задачах. С увеличением числа измерений ее эффективность быстро возрастает. Напомним ее.

В области $G(\mathbf{x})$ строится последовательность сгущающихся вдвое сеток. На самой грубой сетке начальные условия выбираются произвольно; поскольку число узлов этой сетки невелико, то объем вычислений тоже невелик. После установления решение интерполируется на более подробной сетке и выбирается на ней в качестве начальных условий; это в несколько раз уменьшает требуемое для установления число шагов. Затем решение уточняется по способу Рунге с использованием всех сеток.

3. Чебышевский набор шагов. Счет на установление можно проводить с переменным шагом по времени. Для тех задач, в которых известны границы спектра разностных операторов, построены специальные наборы шагов τ_k ($1 \leq k \leq K$), обеспечивающие гораздо более быстрое затухание начальных данных, чем при расчете с постоянным оптимальным шагом.

Чебышевский набор. Пусть разностная схема счета на установление приведена к двуслойной канонической форме:

$$B \frac{y^k - y^{k-1}}{\tau_k} + Ay^{k-1} = \varphi. \quad (28)$$

Будем предполагать, что A и B — самосопряженные положительно определенные операторы, удовлетворяющие неравенствам

$$\gamma_1 B \leq A \leq \gamma_2 B, \quad 0 < \gamma_1 \leq \gamma_2. \quad (29)$$

Затухание начальных данных z^0 определяется однородным уравнением (28), которое можно записать в следующей форме:

$$\zeta^k = (E - \tau_k C) \zeta^{k-1}, \quad (30а)$$

где

$$C = B^{-1/2} A B^{-1/2}, \quad \zeta^k = B^{1/2} z^k. \quad (30б)$$

Отсюда, вытекает, что

$$\zeta^K = P_K(C) \zeta^0, \quad P_K(C) = \prod_{k=1}^K (E - \tau_k C). \quad (31)$$

Для наиболее быстрого затухания начальных данных последовательность шагов τ_k надо выбрать так, чтобы $\|P_K(C)\|$ была минимальна при заданном числе шагов K .

Поскольку A и B — самосопряженные операторы, то оператор C тоже самосопряженный, причем из (29) следует неравенство

$$\gamma_1 E \leq C \leq \gamma_2 E. \quad (32)$$

В этом случае норму операторного многочлена $P_K(C)$ можно оценить по формуле *)

$$\|P_K(C)\| \leq \max_{\gamma_1 \leq \eta \leq \gamma_2} |P_K(\eta)|,$$

где $P_K(\eta)$ — алгебраический многочлен.

Для того чтобы максимум модуля алгебраического многочлена был минимален на отрезке $[\gamma_1, \gamma_2]$, этот многочлен с точностью до множителя должен совпадать с многочленом Чебышева первого рода для этого отрезка **). Используя данные в Приложении корни многочленов Чебышева и учитывая, что корнями $P_K(\eta)$ являются величины $1/\tau_k$, получим чебышевский набор шагов:

$$\tau_k = 2 \left[(\gamma_2 + \gamma_1) + (\gamma_2 - \gamma_1) \cos \frac{\pi(2k-1)}{2K} \right]^{-1}, \quad 1 \leq k \leq K. \quad (33)$$

Определяя множитель, отличающий $P_K(\eta)$ от многочлена Чебышева на отрезке $[\gamma_1, \gamma_2]$, можно найти коэффициент затухания начальных данных после расчета с набором шагов (33):

$$\max_{\gamma_1 \leq \eta \leq \gamma_2} |P_K(\eta)| = \frac{2\rho^K}{1+\rho^{2K}}, \quad \rho = \frac{\sqrt{\gamma_2} - \sqrt{\gamma_1}}{\sqrt{\gamma_2} + \sqrt{\gamma_1}}. \quad (34)$$

Если необходимая точность расчета $\varepsilon \ll 1$, то в оценке (34) можно пренебречь членом ρ^{2K} . Тогда требуемое число шагов равно

$$K(\varepsilon) \approx \frac{1}{\ln(1/\rho)} \ln \frac{2}{\varepsilon}. \quad (35)$$

Заметим, что сначала надо найти требуемое число шагов по формуле (35); только после этого можно вычислить искомый набор шагов (33).

З а м е ч а н и е. В случае области сложной формы или задачи с переменными коэффициентами (2) точные границы энергетической эквивалентности операторов (29) установить обычно не удастся. Приходится оценивать их, занижая γ_1 и завышая γ_2 (в неизвестные числа раз ν_1 и ν_2). Поскольку всегда $\gamma_2 \gg \gamma_1$, то $[\ln(1/\rho)]^{-1} \approx 1/2 \sqrt{\gamma_2/\gamma_1}$. Отсюда видно, что требуемое число шагов (35) возрастает в $\sqrt{\nu_1 \nu_2}$ раз по сравнению со случаем, когда границы спектра известны точно.

*) Доказательство этого неравенства имеется, например, в [15].

**) О многочленах Чебышева и их свойствах см., например, в [9, 24].

Постоянный шаг. Оптимальный постоянный шаг выбирается так, чтобы начальные данные наиболее сильно затухали за один шаг. Там самым, он является частным случаем чебышевского набора, соответствующим $K=1$. Формулы (33), (34) принимают при этом вид

$$\tau_0 = \frac{2}{\gamma_2 + \gamma_1}, \quad \max_{[\gamma_1, \gamma_2]} |P_1(\eta)| = \frac{\gamma_2 - \gamma_1}{\gamma_2 + \gamma_1}. \quad (36)$$

Продольно-поперечная схема (13) или локально-одномерная схема с полусуммой (22) не подходят, строго говоря, под разобранный выше случай, потому что они содержат оператор

$$B_k = \left(E - \frac{1}{2} \tau_k \Lambda_1 \right) \left(E - \frac{1}{2} \tau_k \Lambda_2 \right),$$

явно зависящий от номера шага.

Однако для этих схем в п. 2 были определены оптимальный шаг (19) и соответствующая ему скорость затухания начальных данных. Ограничиваясь задачей Дирихле в p -мерном кубе со стороной a и одинаковым числом узлов N по каждой координате, запишем:

$$\tau_0 = \frac{a^2}{\pi N}, \quad \max |\lambda(\tau_0)| \approx 1 - \frac{\pi p}{N}.$$

Сравнивая эти выражения с (36) и учитывая, что $\gamma_2 \gg \gamma_1$, получим нестрогую, но удовлетворительную оценку границ спектра:

$$\gamma_1 \sim \frac{\pi^2 p}{a^2}, \quad \gamma_2 \sim \frac{2\pi N}{a^2}. \quad (37)$$

Подставляя оценку (37) в (34) и (35), получим, что для рассмотренных схем

$$K \sim \sqrt{\frac{N}{2\pi p}} \ln \frac{2}{\varepsilon}. \quad (38)$$

Таким образом, счет на установление по экономичным схемам с чебышевским набором шагов требует всего $K \sim \sqrt{N}$ шагов, в то время как расчет с постоянным оптимальным шагом требует существенно большего числа шагов: $K(\tau_0) \sim N$.

Используя в разностной схеме переменный оператор B_k , можно найти другие наборы шагов, обеспечивающие еще более быстрое установление. Например, для продольно-поперечной схемы в случае задачи Дирихле (1) в прямоугольнике построен жорданов набор шагов (см. [81]), при котором

$$K = \frac{\ln N}{5} \ln \frac{4}{\varepsilon}.$$

Однако для более сложных задач наборы шагов с подобными характеристиками найти пока не удалось.

Порядок шагов. Чебышевский набор шагов позволяет проводить экономичный расчет на установление даже по явной схеме (11.56) с $\sigma = 0$. Запишем эту схему в виде

$$E \frac{y^k - y^{k-1}}{\tau_k} + \sum_{\alpha=1}^p \Lambda_{\alpha} y^{k-1} = f. \quad (39)$$

Операторы схемы (39) постоянны, и нетрудно показать, что для задачи Дирихле в кубе $\gamma_1 = \pi^2 p/a^2$, $\gamma_2 = 4pN^2/a^2$. Поэтому для расчета по схеме (39) с чебышевским набором шагов требуется число шагов

$$K = \frac{N}{\pi} \ln \frac{2}{\varepsilon}, \quad (40)$$

что по объему вычислений эквивалентно экономичным схемам с постоянным оптимальным шагом.

Заметим, что схема (39) устойчива только при $\tau \leq \tau_0 = h^2/(2p)$. Среди шагов чебышевского набора (33) есть такие, которые больше τ_0 и меньше τ_0 . Большие шаги вызывают рост погрешностей, а малые — затухание. В целом их действие таково, что если выполнить все $K(\varepsilon)$ шагов, то ошибка затухает в ε^{-1} раз.

Слово «если» употреблено не случайно. Нередко ошибки на промежуточных шагах возрастают в $10^{10} - 10^{30}$ раз по сравнению с начальными, выходят за допустимые на ЭВМ пределы, и расчет не удается довести до конца. Поэтому, хотя чебышевский набор шагов для схемы (39) был найден более полувека назад, в практических вычислениях его долго не могли использовать.

Однако если шаги выполнять в определенном порядке, то расчет становится возможным. Идея упорядочения заключается в том, что сразу вслед за шагом, увеличивающим ошибку, надо выполнять шаг, уменьшающий ее. Правило перестановки особенно просто, если число шагов равно 2^r . Тогда надо расположить шаги в естественном порядке и сгруппировать парами: первый — последний, второй — предпоследний и т. д. Затем пары так же группируются в четверки: первая — последняя. Аналогично группируются четверки, восьмерки и т. д. Например, для 16 шагов окончательный порядок такой:

$$1, 16, 8, 9, 4, 13, 5, 12, 2, 15, 7, 10, 3, 14, 6, 11.$$

При использовании упорядоченного чебышевского набора шагов ошибка на отдельных шагах может нарастать, но никогда в ходе расчета не превзойдет начальной ошибки, а в конце расчета будет соответствовать оценке $\|P_K\|$.

§ 2. Вариационные и вариационно-разностные методы

1. Метод Ритца. Вариационные методы применяются к эллиптическим уравнениям в частных производных независимо от числа измерений. Рассмотрим, например, задачу:

$$\operatorname{div} [k(\mathbf{x}) \operatorname{grad} u] - \rho(\mathbf{x}) u(\mathbf{x}) = -f(\mathbf{x}), \quad \mathbf{x} \in G, \quad (41a)$$

$$u_{\Gamma} = \mu(\mathbf{x}). \quad (41б)$$

Дифференциальный оператор $A = -\operatorname{div} [k \operatorname{grad} (\cdot)] + \rho$ является самосопряженным. Поэтому задача (41) эквивалентна задаче на минимум функционала $\Phi[u] = (Au - 2f, u)$, которую при помощи формул векторного анализа можно записать в виде

$$\int_G [k(\mathbf{x}) (\operatorname{grad} u)^2 + \rho(\mathbf{x}) u^2(\mathbf{x}) - 2f(\mathbf{x}) u(\mathbf{x})] d\mathbf{x} = \min, \quad u_{\Gamma} = \mu(\mathbf{x}). \quad (42)$$

Возьмем некоторую функцию $\varphi_0(\mathbf{x})$, удовлетворяющую граничному условию (41б), и полную систему функций $\varphi_l(\mathbf{x})$, $l=1, 2, \dots$, обращающихся в нуль на границе. Будем искать приближенное решение задачи (42) в следующем виде:

$$u(\mathbf{x}) \approx y_n(\mathbf{x}) = \varphi_0(\mathbf{x}) + \sum_{l=1}^n c_l \varphi_l(\mathbf{x}). \quad (43)$$

Подставляя (43) в (42), получим задачу на минимум квадратичной функции неизвестных коэффициентов c_l ; для простоты ограничимся случаем $\varphi_0(\mathbf{x}) \equiv 0$, соответствующим $u_{\Gamma} = 0$:

$$\int_G \left[\sum_{r=1}^n \sum_{l=1}^n c_r c_l (k \operatorname{grad} \varphi_r \operatorname{grad} \varphi_l + \rho \varphi_r \varphi_l) - 2f \sum_{r=1}^n c_r \varphi_r \right] d\mathbf{x} = \min. \quad (44)$$

Приравнивая нулю производные по коэффициентам, получим для определения c_l систему линейных уравнений

$$\sum_{l=1}^n c_l \int_G (k \operatorname{grad} \varphi_r \operatorname{grad} \varphi_l + \rho \varphi_r \varphi_l) d\mathbf{x} = \int_G f \varphi_r d\mathbf{x}, \quad 1 \leq r \leq n. \quad (45)$$

Обоснование сходимости метода Ритца при $n \rightarrow \infty$ рассматривалось в главе VII. При практическом применении метода Ритца успех сильно зависит от выбора системы функций $\varphi_l(\mathbf{x})$. При неудачном выборе этой системы для получения удовлетворительной точности может потребоваться очень много членов ряда (43).

Если область имеет несложную форму, то нередко выбирают систему с разделяющимися переменными; например, для прямоугольника полагают $\varphi_{lm}(\mathbf{x}) = \xi_l(x_1) \eta_m(x_2)$, а для круга $\varphi_{lm}(\mathbf{x}) = \xi_l(r) \eta_m(\theta)$. Отметим, что если в одномерной задаче для полу-

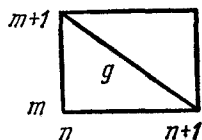
чения удовлетворительной точности требовалось $\sim n$ членов ряда, то в аналогичной p -мерной задаче обычно надо брать $\sim n^p$ членов.

Ограничиваясь малым числом членов, можно легко получить грубую оценку решения.

Замечание 1. Метод Рунта применим к многомерной задаче Штурма — Лиувилля (задаче на собственные значения).

Замечание 2. Если оператор в задаче типа (41) не самосопряженный, то вместо метода Рунта применяют метод Галеркина.

2. Стационарные разностные схемы. Такие схемы можно составлять, непосредственно аппроксимируя производные разностями, или при помощи интегро-интерполяционного метода. Например, для многомерного уравнения $\Delta u = -f(\mathbf{x})$ простейшая разностная замена производных приводит к схеме



$$\sum_{\alpha=1}^p \Lambda_{\alpha} y = -f. \quad (46)$$

Рис. 88.

Составлять разностные схемы можно также вариационными методами. Для этого специальным образом выбирают пробные функции $y_n(\mathbf{x})$, например, считая их сплайнами, построенными по узловым значениям y .

Пример. Рассмотрим решение двумерного уравнения $\Delta u = -f$ на прямоугольной сетке с шагами h_1, h_2 . Эквивалентная задача на минимум в этом случае имеет вид

$$\Phi[u] = \int_G [(\text{grad } u)^2 - 2f(\mathbf{x})u(\mathbf{x})] dx_1 dx_2 = \min \quad (47)$$

(для простоты мы опускаем краевые условия). Разобьем каждую прямоугольную ячейку на две треугольных (рис. 88) и в треугольных ячейках аппроксимируем $u(\mathbf{x})$ линейными функциями; например, в нижнем треугольнике $g(\mathbf{x})$

$$u(\mathbf{x}) \approx y(\mathbf{x}) = y_{nm} + \frac{1}{h_1} (x_1 - x_{1n}) (y_{n+1, m} - y_{nm}) + \frac{1}{h_2} (x_2 - x_{2m}) (y_{n, m+1} - y_{nm}), \quad \mathbf{x} \in g. \quad (48)$$

Совокупность этих функций образует линейный сплайн. Очевидно,

$$(\text{grad } y)^2 = \frac{1}{h_1^2} (y_{n+1, m} - y_{nm})^2 + \frac{1}{h_2^2} (y_{n, m+1} - y_{nm})^2, \quad \mathbf{x} \in g.$$

Аппроксимируя правую часть $f(\mathbf{x})$ в ячейке $g(\mathbf{x})$, например,

константой f_{nm} , легко вычисляем интеграл по этой ячейке:

$$\int_{\xi} [(\text{grad } y)^2 - 2fy] dx = \frac{h_1 h_2}{2} \left[\frac{1}{h_1^2} (y_{n+1, m} - y_{nm})^2 + \right. \\ \left. + \frac{1}{h_2^2} (y_{n, m+1} - y_{nm})^2 - \frac{2}{3} f_{nm} (y_{nm} + y_{n+1, m} + y_{n, m+1}) \right].$$

Аналогично вычисляется интеграл по верхней треугольной ячейке. Суммируя эти интегралы, получим

$$\Phi[y] = \frac{1}{2} h_1 h_2 \sum_{n, m} \left[\frac{1}{h_1^2} (y_{n+1, m} - y_{nm})^2 + \frac{1}{h_1^2} (y_{n+1, m+1} - y_{n, m+1})^2 + \right. \\ \left. + \frac{1}{h_2^2} (y_{n, m+1} - y_{nm})^2 + \frac{1}{h_2^2} (y_{n+1, m+1} - y_{n+1, m})^2 - \frac{2}{3} f_{nm} (y_{nm} + \right. \\ \left. + y_{n+1, m} + y_{n, m+1}) - \frac{2}{3} f_{n+1, m+1} (y_{n+1, m+1} + y_{n+1, m} + y_{n, m+1}) \right] = \\ = \min. \quad (49)$$

Функционал $\Phi[y] = F(y_{nm})$ является квадратичной функцией узловых значений. Приравнявая нулю производные функционала по y_{nm} и учитывая, что эта величина входит в четыре члена двойной суммы (49), получим разностную схему

$$\frac{1}{h_1^2} (y_{n+1, m} - 2y_{nm} + y_{n-1, m}) + \frac{1}{h_2^2} (y_{n, m+1} - 2y_{nm} + y_{n, m-1}) + \\ + \frac{1}{6} (2f_{nm} + f_{n+1, m} + f_{n-1, m} + f_{n, m+1} + f_{n, m-1}) = 0. \quad (50)$$

Это — стационарная схема; легко видеть, что она аппроксимирует непосредственно исходное дифференциальное уравнение.

Замечание. При помощи вариационного метода удобно составлять разностные схемы высокой точности. Для этого решение $u(x)$ и правую часть $f(x)$ аппроксимируют сплайнами более высокого порядка, обычно кубическими (такая аппроксимация обсуждалась в гл. VII, § 4, п. 4).

3. Прямые методы решения. Для стационарных схем типа (47) наиболее сложным является вопрос о фактическом вычислении разностного решения.

В самом деле, p -мерная схема (46) является линейной алгебраической системой с N^p неизвестными (если по каждой координате взято N интервалов). Матрица этой системы в двумерном случае изображена на рис. 89. В общем случае эта матрица ленточная, причем лента слабо заполнена и имеет полуширину N^{p-1} .

Вычисление разностного решения методом исключения Гаусса (который не может использовать слабое заполнение ленты) тре-

бует $\sim N^{3p-2}$ действий, т. е. $\sim N^{2p-2}$ операций на узел сетки *). При большом N это число действий неприемлемо велико; кроме того, лента матрицы не помещается в оперативной памяти ЭВМ. Поэтому прямое решение линейной системы (46) методом Гаусса возможно только в двумерных расчетах, и то при небольшом $N \lesssim 50$.

Замечание. Если строить схемы высокого порядка точности (например, сплайновые) или использовать последовательность сгущающихся сеток (обычно при $N = 4, 8, 16, 32$) с уточнением по способу Рунге, то даже при небольшом числе узлов удастся получить удовлетворительную точность расчета.

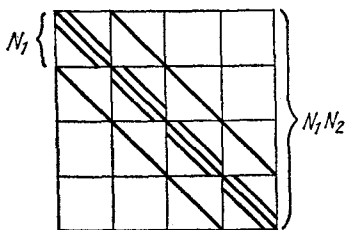


Рис. 89.

Для некоторых важных частных случаев эллиптических задач разработаны очень быстрые прямые методы; перечислим их (они подробно изложены, например, в [3, 6, 31]).

Быстрое преобразование Фурье применимо к задаче Дирихле в прямоугольнике. Оно основано на том, что если число интервалов по каждой переменной N_α разбивается на множители, то вычислять коэффициенты дискретного преобразования Фурье можно не по формулам Бесселя типа (2.44), а по более экономичным рекуррентным формулам. Если $N_\alpha = 2^r \alpha$, то метод является особенно быстрым и требует всего $4 \log_2 N$ действий на каждый узел сетки.

Рассмотрим этот метод сначала на примере одномерной задачи для уравнения с постоянными коэффициентами $u'' - \mu u = -f(x)$ и краевыми условиями первого рода (без ограничения общности их можно взять периодическими). Составим разностную схему на равномерной сетке $\{x_n = nh, 0 \leq n \leq N\}$:

$$\frac{1}{h^2} (y_{n-1} - 2y_n + y_{n+1}) - \mu y_n = -\varphi_n, \quad 1 \leq n \leq N-1, \quad (51)$$

$$y_0 = \varphi_0, \quad y_N = \varphi_0.$$

Будем искать разностное решение в виде разложения Фурье:

$$y_n = \sum_{q=0}^{N-1} a_q \omega^{nq}, \quad \text{где } \omega = \exp(2\pi i/N) \quad (52)$$

*) Аналогичная ситуация возникла в неявной схеме (11.56) для многомерного уравнения теплопроводности.

Подставим разложение (52) в соотношение (51), умножим на $\omega^{-np} = \exp(-2\pi i np/N)$ и просуммируем по n от 0 до $N-1$. Замечая, что

$$\omega^{(n-1)q} - 2\omega^{nq} + \omega^{(n+1)q} = -4\omega^{nq} \sin^2 \frac{\pi q}{N},$$

и учитывая условие ортогональности гармоник (см. гл. II, § 2, п. 4), найдем

$$a_p = b_p / \left(\frac{4}{h^2} \sin^2 \frac{\pi p}{N} + \mu \right), \quad (53)$$

где

$$b_p = \frac{1}{N} \sum_{n=0}^{N-1} \varphi_n \omega^{-np}, \quad 0 \leq p \leq N-1, \quad (54)$$

являются дискретными коэффициентами Фурье правой части уравнения. Формулы (53), (54) позволяют найти искомое разностное решение.

Однако эти формулы неэкономичны. Необходимо вычислить N коэффициентов b_p , причем нахождение каждого коэффициента по формуле (54) требует примерно $2N$ операций. Следовательно, задача (51) решается за $2N^2$ операций, т. е. много медленнее, чем в методе прогонки.

Если число интервалов сетки составное, $N = KL$, то формулу (54) можно преобразовать так, что требуемое количество операций уменьшится. Представим индексы n и p в следующем виде:

$$n = l_1 + Ll_2, \quad 0 \leq l_1 \leq L-1, \quad 0 \leq l_2 \leq K-1,$$

$$p = p_1 + Kp_2, \quad 0 \leq p_1 \leq K-1, \quad 0 \leq p_2 \leq L-1.$$

Запишем формулу (54) в виде двойной суммы:

$$b_p = \frac{1}{KL} \sum_{l_1=0}^{L-1} \sum_{l_2=0}^{K-1} \varphi_{l_1 + Ll_2} \omega^{-p_1 l_1 - L p_1 l_2 - K l_1 p_2 - L K l_2 p_2}.$$

Отбросим в показателе степени последнее слагаемое, ибо $\omega^{LK} = 1$, и получим следующее выражение коэффициентов Фурье:

$$b(p) \equiv b_p = \frac{1}{L} \sum_{l_1=0}^{L-1} b(l_1, p_1) \omega^{-p_1 l_1}, \quad 0 \leq p \equiv p_1 + Kp_2 \leq N-1, \quad (55)$$

где

$$b(l_1, p_1) = \frac{1}{K} \sum_{l_2=0}^{K-1} \varphi_{l_1 + Ll_2} \omega^{-L p_1 l_2}, \quad 0 \leq l_1 \leq L-1, \quad 0 \leq p_1 \leq K-1. \quad (56)$$

Вычисление N коэффициентов $b(p)$ по формуле (55) требует $2NL$ операций; вычисление $LK = N$ вспомогательных коэффициентов $b(l_1, p_1)$ по формуле (56) производится еще за $2NK$ операций. Следовательно, число операций, необходимое для нахождения коэффициентов Фурье по формулам (55), (56), равно $2N(L+K)$; оно существенно меньше, чем $2N^2$ (например, при $K=L=\sqrt{N}$ меньше в $\sqrt{N}/2$ раз).

Если K в свою очередь разбивается на множители, то формулу (56) следует преобразовать аналогичным образом. Это позволяет еще уменьшить объем вычислений.

Приведем без вывода рекуррентные формулы вычисления коэффициентов Фурье для случая $N = L^r$:

$$b(p) = \frac{1}{L} \sum_{l_1=0}^{L-1} b(l_1, p_1) \omega^{-pl_1},$$

$$\begin{aligned} b(l_1, l_2, \dots, l_k, p_k) &= \\ &= \frac{1}{L} \sum_{l_{k+1}=0}^{L-1} b(l_1, l_2, \dots, l_{k+1}, p_{k+1}) \omega^{-L^k l_{k+1} p_k}, \\ &1 \leq k \leq r-2, \end{aligned} \quad (57)$$

$$\begin{aligned} b(l_1, l_2, \dots, l_{r-1}, p_{r-1}) &= \\ &= \frac{1}{L} \sum_{l_r=0}^{L-1} \Phi_{l_1 + L l_2 + L^2 l_3 + \dots + L^{r-1} l_r} \omega^{-L^{r-1} l_r p_{r-1}}, \end{aligned}$$

причем

$$0 \leq l_k \leq L-1, \quad 0 \leq p_k \leq L^{r-k} - 1.$$

Число вспомогательных коэффициентов k -го ранга $b(l_1, \dots, l_k, p_k)$ равно N , поэтому для вычисления коэффициентов всех рангов по формулам (57) требуется около $2NLr$ операций.

Если учесть, что $L = N^{1/r}$, то нетрудно найти оптимальное число сомножителей $r_{\text{опт}} \approx \ln N$ и оптимальное значение $L_{\text{опт}} \approx e \approx 3$. Но для программирования считается более удобным, если $N = 2^r$ и $L = 2$; в последнем случае требуемое число операций равно $4N \log_2 N$, что мало отличается от оптимального случая и почти не уступает по скорости методу прогонки.

Обобщение этого метода на случай многих измерений очевидно. Пусть, например, для уравнения с постоянными коэффициентами

$$u_{x_1 x_1} + u_{x_2 x_2} - \mu u = -f(x_1, x_2)$$

поставлена первая краевая задача в прямоугольной области. Введем равномерную сетку $\{x_{1n} = nh_1, x_{2m} = mh_2, 0 \leq n \leq N, 0 \leq$

$\leq m \leq M\}$ и составим разностную схему

$$\frac{1}{h_1^2} (y_{n-1, m} - 2y_{nm} + y_{n+1, m}) + \frac{1}{h_2^2} (y_{n, m-1} - 2y_{nm} + y_{n, m+1}) - \mu y_{nm} = -\varphi_{nm}. \quad (58)$$

Будем искать разностное решение в виде разложения Фурье

$$y_{nm} = \sum_{p=0}^{N-1} \sum_{q=0}^{M-1} a_{pq} \omega_1^{np} \omega_2^{mq}, \quad (59)$$

$$\omega_1 = \exp(2\pi i/N), \quad \omega_2 = \exp(2\pi i/M).$$

Аналогично одномерному случаю, получим следующие выражения для коэффициентов Фурье:

$$a_{pq} = b_{pq} / \left(\frac{4}{h_1^2} \sin^2 \frac{\pi p}{N} + \frac{4}{h_2^2} \sin^2 \frac{\pi q}{M} + \mu \right), \quad (60)$$

где

$$b_{pq} = \frac{1}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \varphi_{nm} \omega_1^{-np} \omega_2^{-mq}. \quad (61)$$

Запишем последнюю формулу в следующем виде:

$$b_{pq} = \frac{1}{N} \sum_{n=0}^{N-1} \beta_{nq} \omega_1^{-np}, \quad (62)$$

$$\beta_{nq} = \frac{1}{M} \sum_{m=0}^{M-1} \varphi_{nm} \omega_2^{-mq}.$$

Каждая сумма в формулах (62) имеет тот же вид, что и в формуле (54). Поэтому, если N и M разлагаются на множители, то каждую сумму можно вычислить по рекуррентным формулам типа (57). Если при этом $N = L_1^{r_1}$ и $M = L_2^{r_2}$, то число операций на каждый узел сетки, аналогично одномерному случаю, есть $O(r_1 L_1 + r_2 L_2) = O(\log(NM))$. Следовательно, быстрое преобразование Фурье даже в многомерном случае по экономичности мало уступает самому быстрому одномерному методу — прогонке.

Метод декомпозиции, или нечетно-четной редукции, применим для той же задачи, что и быстрое преобразование Фурье. Он использует исключение всех нечетных точек из системы уравнений типа (46). При $N_\alpha = 2^{r_\alpha}$ исключение выполняет-

ся рекуррентно и число действий на узел сетки составляет $O(\log_2 N)$.

Матричная прогонка применима даже для случая областей сложной формы. Число действий на узел сетки в этом методе есть $O(N^2)$. Но если требуется решить на данной сетке большую серию задач с различными правыми частями и граничными значениями, то, сохраняя и используя результаты промежуточных вычислений, можно сократить это число действий до $O(N)$.

Быстрые прямые методы обобщены в настоящее время на задачи в круге и области ступенчатой формы (в этих случаях их скорость падает). Однако для областей произвольной формы, а также для уравнения достаточно общего вида (2) удовлетворительных прямых методов пока не найдено.

4. Итерационные методы. В случае сложных задач неэволюционные разностные схемы $Ay = -f$ решают итерационными методами. Простейшим из них является метод Якоби (5.51), относящийся к методам последовательного приближения. Для двумерного уравнения (46) он имеет вид

$$\left(\frac{2}{h_1^2} + \frac{2}{h_2^2}\right) y_{nm}^k = \frac{1}{h_1^2} (y_{n+1, m}^{k-1} + y_{n-1, m}^{k-1}) + \frac{1}{h_2^2} (y_{n, m-1}^{k-1} + y_{n, m+1}^{k-1}) + f_{nm},$$

где k — номер итерации. Это выражение можно формально переписать следующим образом:

$$E \frac{y^k - y^{k-1}}{\tau} + \sum_{\alpha=1}^2 \Lambda_{\alpha} y^{k-1} = -f, \quad \tau = \left(\frac{2}{h_1^2} + \frac{2}{h_2^2}\right)^{-1}.$$

Большинство итерационных методов можно символически записать в аналогичной форме:

$$B_k \frac{y^k - y^{k-1}}{\tau_k} + Ay^{k-1} = -f; \quad (63)$$

если B и τ не зависят от номера итерации k , то процесс называют *стационарным*.

Итерационный процесс (63) можно рассматривать как разностную схему расчета некоторой эволюционной задачи. Эту задачу можно найти, определяя нулевое или первое дифференциальное приближение разностной схемы (63). Физический смысл найденной задачи не имеет значения; важно только, чтобы она соответствовала диссипативному процессу, т. е. обеспечивала бы установление стационарного решения. Тем самым, *несущественно, что именно аппроксимирует оператор B* ; он должен только: а) обеспечивать возможно более быстрое затухание начальных данных и б) легко обращаться, чтобы решение y^k вычислялось за малое число действий.

Продольно-поперечная и локально-одномерная схемы, которые можно формально рассматривать как итерационные процессы (63) для решения системы $Ay = -f$, удовлетворяют этим двум требованиям. Однако этим требованиям удовлетворяют также некоторые схемы, невыгодные для расчета параболических задач. Одной из таких схем является

Попеременно-треугольная схема, которую мы рассмотрим на примере двумерного уравнения

$$k_1 u_{x_1 x_1} + k_2 u_{x_2 x_2} = -f(x_1, x_2).$$

Запишем вспомогательное параболическое уравнение:

$$v_t = k_1 v_{x_1 x_1} + k_2 v_{x_2 x_2} + f(x_1, x_2).$$

Выберем шаблон, показанный на рис. 90, и составим на нем чисто неявную разностную схему

$$\frac{1}{\tau} (\hat{y} - y) = (\Lambda_1 + \Lambda_2) \hat{y} + \varphi,$$

которую можно переписать в канонической форме:

$$(E - \tau\Lambda_1 - \tau\Lambda_2) \frac{\hat{y} - y}{\tau} - (\Lambda_1 + \Lambda_2) y = \varphi. \quad (64)$$

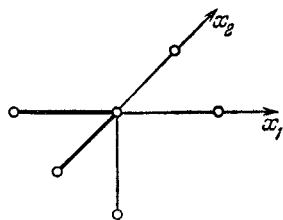


Рис. 90.

Как отмечалось выше, эта схема неэкономична, поскольку обращение оператора $E - \tau\Lambda_1 - \tau\Lambda_2$ требует в общем случае большого числа операций на каждый узел сетки.

Введем *треугольные* операторы

$$\begin{aligned} R_1 y_{nm} &= \frac{k_1}{h_1^2} (y_{nm} - y_{n-1, m}) + \frac{k_2}{h_2^2} (y_{nm} - y_{n, m-1}), \\ R_2 y_{nm} &= \frac{k_1}{h_1^2} (y_{nm} - y_{n+1, m}) + \frac{k_2}{h_2^2} (y_{nm} - y_{n, m+1}), \end{aligned} \quad (65)$$

определенные на треугольных шаблонах (шаблон для R_1 показан на рис. 90 жирными линиями). Нетрудно заметить, что $\Lambda_1 + \Lambda_2 = = - (R_1 + R_2)$, что позволяет записать схему (64) следующим образом:

$$(E + \tau R_1 + \tau R_2) \frac{\hat{y} - y}{\tau} - (\Lambda_1 + \Lambda_2) y = \varphi. \quad (66)$$

Слегка изменим схему (66), добавляя в левую часть член $\tau^2 R_1 R_2 (\hat{y} - y)/\tau$, имеющий порядок малости $O(\tau^2)$. Возникающий при этом оператор $E + \tau R_1 + \tau R_2 + \tau^2 R_1 R_2$ факторизуется, т. е. представляется в виде произведения операторов $E + \tau R_1$ и $E + \tau R_2$.

Полученную схему называют *попеременно-треугольной*:

$$(E + \tau R_1)(E + \tau R_2) \frac{\hat{y} - y}{\tau} - (\Lambda_1 + \Lambda_2) y = \varphi. \quad (67)$$

Операторы $E + \tau R_\alpha$ легко обращаются, так что алгоритм вычисления разностного решения в этой схеме несложен и требует небольшого числа операций на каждый узел сетки. В самом деле, такие операторы уже встречались при составлении схемы бегущего счета (10.29) для многомерного уравнения переноса; организация вычислений в этом случае была подробно разобрана в гл. X, § 1, п. 4. Схема (67) также решается посредством бегущего счета. На каждом слое сначала обращают оператор $E + \tau R_1$; вычисления при этом начинают с узла (x_{10}, x_{20}) и ведут, например, по направлениям x_1 , доходя в конечном итоге до узла (x_{1N}, x_{2M}) . Затем обращают оператор $E + \tau R_2$, начиная вычисления с узла (x_{1N}, x_{2M}) и ведя их в обратном порядке.

Попеременно-треугольная схема естественно переносится на случай любого числа измерений. Она легко обобщается на дифференциальные уравнения с переменными или разрывными коэффициентами и области $G(\mathbf{x})$ сложной формы. При этом схема для исходной задачи $Ay = -f$ записывается в виде

$$B_k = (D + \tau_k R_1) D^{-1} (D + \tau_k R_2), \quad B_k = \frac{y^k - y^{k-1}}{\tau_k} + Ay^{k-1} = f, \quad (68)$$

где $D = D^H > 0$ — диагональный оператор, выбираемый так, чтобы возможно сильнее уменьшить отношение γ_2/γ_1 границ эквивалентности (29) операторов A и B_k , а треугольные операторы R_α выбраны так, чтобы выполнялось $R_1 + R_2 = A$, $R_1 = R_2^H$ (нетрудно заметить, что в схеме (67) эти условия выполнены, причем $D = E$). Если используется чебышевский набор шагов, то процесс (68) сходится за $K = O\left(\sqrt{N} \ln \frac{1}{\varepsilon}\right)$ итераций.

Градиентные методы. Можно заменить линейную задачу $Ay = -f$ задачей на минимум квадратичной функции $F(y)$. Если матрица A положительно определенная, то удобно взять задачу

$$F(y) \equiv (Ay, y) + 2(f, y) = \min. \quad (69)$$

Для произвольной матрицы A (которая встречается в задачах со смешанными производными) можно положить

$$F(y) \equiv (Ay + f, Ay + f) = \min. \quad (70)$$

Задачу на минимум можно решать методом наискорейшего спуска, что для случая (69) выполняется по формулам (6.22) — (6.26).

Скорость сходимости метода наискорейшего спуска, согласно оценке (6.27), такая же, как и у экономических схем с постоянным оптимальным шагом, т. е. $K(\varepsilon) = O(N \ln(1/\varepsilon))$. Она меньше, чем у схем с чебышевским набором шагов. Достоинством метода является то, что для его применения не надо знать границы спектра оператора A .

ЗАДАЧИ

1. Найти время, необходимое для установления стационарного режима в эволюционной задаче (10), и исследовать характер установления.

2. Найти оптимальный шаг для счета на установление по локально-одномерной схеме типа (22) в случае задачи Дирихле (1) в трехмерном параллелепипеде.

3. Найти оптимальный шаг и необходимое число шагов $K(\varepsilon)$ для счета на установление по явной схеме (39) в случае задачи Дирихле в p -мерном параллелепипеде, когда сетки равномерны, а число узлов по каждой переменной N_α свое.

4. Для условий задачи 3 построить упорядоченный чебышевский набор шагов при $K=64$.

5. Обосновать критерий установления (276).

6. Для решения задачи (41) методом Рунге написать аналог системы (45) при $\varphi_0(x) \neq 0$.

7. Составить вариационным методом разностную схему, аналогичную (50), используя для $f(x)$ сплайновую аппроксимацию типа (48).

8. Составить формулы наискорейшего спуска для задачи (70).

9. Доказать справедливость рекуррентных формул (57). Указание: полагая последовательно $N = LN_1$, $N_1 = LN_2$ и т. д., использовать для индексов следующую замену переменных:

$$\begin{aligned} p &= p_1 + N_1 p'_1, & p_1 &= p_2 + N_2 p'_2, & p_2 &= p_3 + N_3 p'_3, & \dots, \\ n &= l_1 + Ll'_2, & l'_2 &= l_2 + Ll'_3, & l'_3 &= l_3 + Ll'_4, & \dots \end{aligned}$$