ЧИСЛЕННЫЕ МЕТОДЫ РЕШЕНИЯ ЗАДАЧ ТЕОРИИ УПРАВЛЕНИЯ

Е. С. Тверская

МГТУ им. Н.Э. Баумана
Москва
E-mail: e_tverskaya@bmstu.ru

1. ПАРАБОЛИЧЕСКИЕ УРАВНЕНИЯ

К параболическим уравнениям относятся уравнения теплопроводности, диффузии и т.д. Типичной постановкой задачи является задача теплопродности с граничными условиями первого рода.

$$\begin{aligned}
&u_t(x,t) = a^2 u_{xx}(x,t) + f(x,t),\ 0 < x < l,\ t > 0\\
&u(x,0) = \mu(x),\ 0 \leqslant x \leqslant l;\\
&u(0,t) = \mu_1(t),\ u(l,t) = \mu_2(t),\ 0 \leqslant t \leqslant T.
\end{aligned}$$

В данном случае можно рассмотреть линейное пространство дважды непрерывно дифференцируемых функций $u(x,t)$ и в нем линейный оператор L, который каждой функции $u(x,t)$ ставит в соответствие упорядоченный набор

$$L(u) = \left(u_t - a^2 u_{xx}; u(x,0); u(0,t); u(l,t)\right),$$

являющийся элементом соответствующего линейного пространства. Тогда краевую задачу можно записать в виде $Lu = g$, где f — это упорядоченный набор

$$g = (f(x,t); \mu(x); \mu_1(t); \mu_2(t)).$$

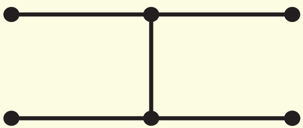
Рассматриваемые линейные пространства бесконечномерны. При дискретизации операторное уравнение $Lu = g$ заменяется серией уравнений $\hat{L}_N \hat{u}_N = \hat{g}_N$, в котором $\hat{u}_N$ — ***сеточная функция***, т.е. некоторая функция, определенная в узлах сетки $\omega_h + \gamma_h$, $\hat{L}_N$ — векторная функция, описывающая левые части сеточных уравнений, а $\hat{g}_N$ — вектор, описывающий правые части сеточных уравнений.

1.1. Семейство неявных схем

Рассмотрим простейшие разностные схемы для уравнения теплопроводности. Перепишем уравнение теплопроводности в наших обозначениях:

$$u_t(x,t) = k\,u_{xx}(x,t) + f(x,t),\ 0 < x < a,\ 0 < t \leqslant T$$

Возьмем в области $G = \{0, a\} \times \{0, T\}$ прямоугольную сетку. Пусть рассматриваемая сетка будет равномерной с шагами h и τ. Выберем шеститочечный шаблон, изображенный на рисунке.



Составим для данного шаблона двуслойную схему (данная схема была задана на дом).

$$\frac{1}{\tau}\left(\hat{y}_n - y_n\right) = \frac{k\sigma}{h^2}\left(\hat{y}_{n-1} - 2\hat{y}_n + \hat{y}_{n+1}\right) + \frac{k(1-\sigma)}{h^2}\left(y_{n-1} - 2y_n + y_{n+1}\right) + \varphi_n,\ 1 \leqslant n \leqslant N-1. \tag{1.1}$$

Краевые условия в рассматриваемой задаче аппроксимируются следующим образом

$$\hat{y}_0 = \mu_1\left(t_{m+1}\right),\ \hat{y}_N = \mu_2\left(t_{m+1}\right)$$

Замечание 1.1. Как правило, значение φ_n вычисляют как $f(x_n, t + \tau/2)$.

Перегруппировав слагаемые получим:

$$\hat{y}_{n-1} + \left(2 + \frac{h^2}{k\tau\sigma}\right)\hat{y}_n + \hat{y}_{n+1} = \left(-\frac{h^2}{k\tau\sigma} + \frac{2(1-\sigma)}{\sigma}\right)y_n - \frac{(1-\sigma)}{\sigma}\left(y_{n-1} + y_{n+1}\right) - \frac{h^2}{k\sigma}\varphi_n;$$

$$\hat{y}_0 = \mu_1\left(t+\tau\right),\ \hat{y}_N = \mu_2\left(t+\tau\right).$$

Полученная система имеет трехдиагональную матрицу и может быть решена методом прогонки. Достаточные условия реализуемости метода прогонки выполнены, так как матрица системы имеет строгое диагональное преобладание.

Замечание 1.2.

- ✓ Если $\sigma = 0$, то неявная рассматриваемая схема (1.1) становится явной. Схема использует только четыре точки шаблона. Разностное решение в этом случае легко вычисляется. Его существование и единственность очевидны.

- ✓ При $\sigma = 1$ схема называется **чисто неявной** и также использует только четыре точки шаблона.

- ✓ При $\sigma = 1/2$ схему называют **схемой с полусуммой** или **симметричной**.

1.1.1. Аппроксимация

Для вычисления порядка аппроксимации разложим решение в узлах шаблона по формуле Тейлора в точке $(x_n, t + \tau/2)$.

Задача 1 *Показать, что при $\sigma \neq 1/2$ схема имее аппроксимацию $O(\tau + h^2)$. Если $\sigma = 1/2$, то схема имеет аппроксимацию $O(\tau^2 + h^2)$*

Решить данную задачу самостоятельно. Для этого необходимо вычислить невязку.

Задача 2 *При* $k = \mathrm{const}$ *построить схему повышенной точности.*

Для этого необходимо дважды продифференцировать уравнение теплопроводности по пространственной переменной:

$$u_{txx} = k u_{xxxx} + f_{xx}.$$

Преобразовать невязку, записав в разложении по формуле Тейлора большее число членов и подставив полученное выражение для u_{txx}. После этого, положив

$$\sigma = \frac{1}{2} - \frac{h^2}{12k\tau}, \quad \varphi_n = \left(f + \frac{h^2}{12} f_{xx}\right)_{x_n, t+\tau/2}$$

Решить данную задачу самостоятельно.

1.1.2. Устойчивость

Исследуем рассматриваемую схему на устойчивость. Для этого рассмотри еще один ранее не описанный способ, а именно:

Метод разделения переменных.

Для поставленной нами задачи требование устойчивости налагает ограничение на параметры сетки h и τ. Рассмотрим смешанную разностную схему, рассматривая явную и неявную схемы как частный случай смешанной. Основной вывод здесь состоит в том, что смешанная схема, определяемая параметром σ, является устойчивой, если

$$\left(\frac{1}{2} - \sigma\right) \max_{x \in [0, l]} K(u(x)) \frac{4\tau}{h^2} \leqslant 1.$$

Отсюда, в частности, следует, что смешанная разностная схема при $\alpha \geqslant 1/2$, в том числе неявная схема, устойчивы при любых соотношениях h и τ. Смешанная разностная схема при $\alpha < 1/2$, в том числе явная схема, устойчивы лишь при определенных сочетаниях параметров h и τ.

В записанном неравенстве коэффициент k не является постоянным, т.е. в общем случае задача выглядит следующим образом:

$$\begin{aligned}
&c\rho\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\Bigl(K(u(x))\frac{\partial u}{\partial x}\Bigr) + f(x,t), \quad x \in (0,l), \quad t > 0;\\
&u(x,0) = \psi(x);\\
&u(0,t) = \mu(t);\\
&K(u)\frac{\partial u}{\partial x}\Big|_{x=l} = W(t).
\end{aligned}$$

Чтобы упростить выкладки, рассмотрим частный случай краевой задачи, когда коэффициент теплопроводности не зависит от температуры, т.е. $K(u) \equiv K$. В этом случае введем оператор

$$(\Lambda \mathbf{g})_n = \frac{K}{h^2}\bigl(g_{n-1} - 2g_n + g_{n+1}\bigr), \quad n = \overline{1, N-1}.$$

И разностная схема в этом случае определяется операторным уравнением

$$\frac{\hat{y}_n - y_n}{\tau} = \sigma\Lambda\hat{y}_n + (1-\sigma)\Lambda y_n + \varphi_n, \tag{1.2}$$

граничными условиями

$$y_0^m = \mu^m, \quad K\frac{y_N^m - y_{N-1}^m}{h} = W^m, \quad m = \overline{1, M}, \tag{1.3}$$

и начальным условием

$$y_n^0 = \psi_n, \quad 0 \leqslant n \leqslant N.$$

Эту задачу удобно свести к случаю однородных граничных условий. Такое преобразование аналогично преобразованию непрерывной задачи: достаточно придумать функцию y_0, удовлетворяющую поставленным граничным условиям и провести замену $y = v + y_0$. Тогда функция v будет удовлетворять однородным граничным условиям, а в операторном уравнении появится дополнительное слагаемое, соответствующее функции y_0. В нашем случае пусть z удовлетворяет граничным условиям (1.3). В операторное уравнение подставим $y = v + z$. Тогда

$$\frac{\hat{v}_n - v_n}{\tau} = \sigma \Lambda \hat{v}_n + (1-\sigma)\Lambda v_n + \mathbf{g}_n,$$

где

$$\mathbf{g}_n = \sigma \Lambda \hat{z}_n + (1-\sigma)\Lambda z_n - \frac{\hat{z}_n - z_n}{\tau} + \varphi_n.$$

Функцию z легко выбрать. Например, можно положить $z_n^m = 0$ при $1 \leqslant n \leqslant N-1$, $z_0^m = \mu^m$ и $z_N^m = hW^m/K$. Множество одномерных сеточных функций, удовлетворяющих однородным граничным условиям, т.е. функций $\mathbf{g} \in \Omega_h$, для которых $g_0 = 0$ и $g_{N-1} = g_N$, есть подпространство в Ω_h. Поскольку такие функции однозначно определяются своими значениями во внутренних узлах, это подпространство можно отождествить с Ω_h^0, где Ω_h — линейное пространство одномерных сеточных функций, принимающих значения в точках x_n, $n = \overline{0, N}$. Ω_h^0 — линейное пространство сеточных функций, которые определены во внутренних узлах пространственной сетки x_n, $n = \overline{1, N-1}$.

Для оценки близости сеточных функций в Ω_h^0 необходимо в этом линейном пространстве ввести норму. Поскольку оно конечномерное, выбор нормы не является существенным и может проводиться из соображений удобства. В Ω_h^0 можно ввести скалярное произведение

$$(u,v) = \sum_{i=1}^{N-1} u_i v_i h$$

и норму $||u|| = \sqrt{u}$, индуцированную скалярным произведением. Мы также будем оперировать нормой

$$||u||_c = \max\{|u_1|; |u_2|; \ldots; |u_{N-1}|\}.$$

Ошибка, получаемая при вычислении $(m+1)$-го слоя формируется ошибками, входящими в значения функций v_n и $\mathbf{g}_n$. Будем считать, что эти функции имеют вид $v_n + \delta v_n$ и $\mathbf{g}_n + \delta \mathbf{g}_n$, где символ δ указывает на погрешность функции. Тогда основное уравнение будет иметь вид

$$\frac{(\hat{v}_n + \delta\hat{v}_n) - (v_n + \delta v_n)}{\tau} = \sigma\Lambda(\hat{v}_n + \delta\hat{v}_n) + (1-\sigma)\Lambda(v_n + \delta v_n) + \mathbf{g}_n + \delta\mathbf{g}_n.$$

Считая, что "истинные" функции v_n, $\hat{v}_n$ и $\mathbf{g}_n$ также связаны операторным уравнением, заключаем, что

$$\frac{\delta\hat{v}_n - \delta v_n}{\tau} = \sigma\Lambda(\delta\hat{v}_n) + (1-\sigma)\Lambda(\delta v_n) + \delta\mathbf{g}_n,$$

т.е. погрешности связаны тем же уравнением. Из последнего уравнения находим

$$\delta\hat{v}_n - \tau\sigma\Lambda(\delta\hat{v}_n) = \delta v_n + \tau(1-\sigma)\Lambda(\delta v_n) + \tau\delta\mathbf{g}_n,$$

или

$$(E - \tau\sigma\Lambda)\delta\hat{v}_n = (E + \tau(1-\sigma)\Lambda)\delta v_n + \tau\delta\mathbf{g}_n.$$

Если оператор $E - \tau\sigma\Lambda$ имеет обратный, то

$$\delta\hat{v}_n = (E - \tau\sigma\Lambda)^{-1}(E + \tau(1-\sigma)\Lambda)\delta v_n + \tau(E - \tau\sigma\Lambda)^{-1}\delta\mathbf{g}_n. \tag{1.4}$$

Как мы видим, в погрешность $\delta\hat{v}_n$ входят две компоненты. Первая связана с ошибкой предыдущего слоя, а вторая — с ошибками в граничных условиях. Отметим, что оператор Λ является самосопряженным, так как в стандартном базисе (т.е. базисе из сеточных функций $\mathbf{e}^k$, для которых $e_i^k = \delta_i^k$, $i, k = \overline{1, N-1}$, где δ_i^k — символ Кронекера) матрица этого оператора симметрическая:

$$[\Lambda]_e = \frac{K}{h^2}\begin{pmatrix} -2 & 1 & 0 & \ldots & 0 & 0 \\ 1 & -2 & 1 & \ldots & 0 & 0 \\ 0 & 1 & -2 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 1 & -1 \end{pmatrix}$$

Такой оператор приводится к диагональному виду, причем элементы соответствующей матрицы — собственные числа — могут быть найдены как решения характеристического уравнения $\det(\Lambda - \mu E) = 0$. Операторы

$$L = (E - \tau\sigma\Lambda)^{-1}(E + \tau(1-\sigma)\Lambda), \qquad F = (E - \tau\sigma\Lambda)^{-1}$$

также являются самосопряженными.

Характеристическое уравнение для L имеет вид

$$\det(L-\mu E)=\det\Big((E-\tau\sigma\Lambda)^{-1}(E+\tau(1-\sigma)\Lambda)-\mu E\Big)=$$
$$=\det(E-\tau\sigma\Lambda)^{-1}\cdot\det\Big((E+\tau(1-\sigma)\Lambda)-\mu(E-\tau\sigma\Lambda)\Big)=$$
$$=\det(E-\tau\sigma\Lambda)^{-1}\cdot\det\Big(\tau(1-\sigma+\sigma\mu)\Lambda-(\mu-1)E\Big)=0.$$

Следовательно, если оператор $E-\tau\sigma\Lambda$ обратим, то уравнение $\det(L-\mu E)=0$ равносильно уравнению

$$\det\left(\Lambda-\frac{\mu-1}{\tau(1-\sigma+\sigma\mu)}E\right)=0.$$

откуда следует, что собственные числа μ_i оператора L связаны с собственными числами λ_i оператора Λ соотношениями

$$\frac{\mu_i-1}{\tau(1-\sigma+\sigma\mu_i)}=\lambda_i.$$

или

$$\mu_i=1+\frac{\tau\lambda_i}{1-\sigma\tau\lambda_i}.$$

Характеристическое уравнение для оператора F имеет вид

$$\det(F-\nu E)=\det\Big((E-\tau\sigma\Lambda)^{-1}-\nu E\Big)=\det(E-\tau\sigma\Lambda)^{-1}\cdot\det(\tau\sigma\nu\Lambda-(\nu-1)E)=0,$$

откуда

$$\det\Big(\Lambda-\frac{\nu-1}{\tau\sigma\nu}E\Big)=0.$$

Следовательно, собственные числа ν_i оператора F можно найти через собственные числа λ_i оператора Λ по формулам

$$\nu_i = \frac{1}{1-\tau\sigma\lambda_i}.$$

Найдем собственные числа оператора Λ. Для этого необходимо найти все нетривиальные сеточные функции v из пространства Ω_h^0, удовлетворяющие операторному уравнению $\Lambda v = \lambda v$. Оператор Λ на сеточных функциях является аналогом оператора $\dfrac{d^2}{dx^2}$ на непрерывных функциях. Поэтому и решения операторного уравнения можно найти по аналогии. Ищем решения указанного операторного уравнения в виде

$$v_n = A\sin\omega n + B\cos\omega n, \quad n=\overline{1, N-1}.$$

Используя вид оператора Λ, находим

$$(\Lambda v)_n = \frac{K}{h^2}(v_{n-1} - 2v_n + v_{n+1}) = \frac{K}{h^2}(2\cos\omega - 2)(A\sin\omega n + B\cos\omega n).$$

самостоятельно подставить v_n и получить правую часть равенства

Нам остается найти среди этих функций те, которые попадают в Ω_h^0, т.е. удовлетворяют однородным граничным условиям $v_0 = 0$ и $v_N = v_{N-1}$. Из первого равенства следует, что $B = 0$, а из второго заключаем, что $\sin\omega N = \sin\omega(N-1)$, или $2\sin\dfrac{\omega}{2}\cos\omega\left(N - \dfrac{1}{2}\right) = 0$. Решениями полученного уравнения будет последовательность

$$\omega_k = \frac{(2k-1)\pi}{2N-1}, \quad k\in\mathbb{Z}.$$

Они определяют $N-1$ независимых сеточных функций v_n в соответствии с формулой

$$v_{nk} = \sin\frac{(2k-1)\pi n}{2N-1}, \quad n, k = \overline{1, N{-}1}.$$

Собственными числами найденных собственных фукнций будут числа

$$\lambda_k = \frac{K}{h^2}\Bigl(2\cos\frac{(2k-1)\pi}{2N-1} - 2\Bigr) = -\frac{4K}{h^2}\sin^2\frac{(2k-1)\pi}{2(2N-1)}, \quad k = \overline{1, N{-}1}$$

В частности, получаем $-4K/h^2 \leqslant \lambda_k \leqslant 0$, а максимальное по модулю собственное число равно

$$\lambda_{\max} = -\frac{4K}{h^2}\cos^2\frac{\pi}{2N-1}.$$

Оператор $E-\tau\sigma\Lambda$ обратим, так как уравнение $(E-\tau\sigma\Lambda)v = 0$ имеет только тривиальное решение. Действительно, это уравнение равносильно уравнению $\Lambda v = (\tau\sigma)^{-1}v$, но линейный оператор Λ не имеет положительных собственных чисел, а значит, и $(\tau\sigma)^{-1}$ не является собственным числом Λ.

Поскольку оператор $E-\tau\sigma\Lambda$ обратим, все ранее выписанные равенства, базирующиеся на предположении обратимости этого оператора, корректны.

Перейдем к оценке ошибки $\hat v_n$. Из равенства (1.4) получаем

$$|\delta\hat v_n| \leqslant |L(\delta v_n)| + \tau|F(\delta\mathbf{g}_n)| \leqslant ||L||\,|\delta v_n| + \tau||F||\,|\delta\mathbf{g}_n|. \tag{1.5}$$

Евклидовы нормы операторов L и F легко находятся через их собственные числа (собственно , именно для этого и определялся их спектр):

$$||L|| = \max|\mu_i| = \max\left\{1 - \frac{\dfrac{4K\tau}{h^2}\sin^2\dfrac{\pi}{2(2N-1)}}{1+\sigma\dfrac{4K\tau}{h^2}\sin^2\dfrac{\pi}{2(2N-1)}}; \left|1 - \frac{\dfrac{4K\tau}{h^2}\cos^2\dfrac{\pi}{2N-1}}{1+\sigma\dfrac{4K\tau}{h^2}\cos^2\dfrac{\pi}{2N-1}}\right|\right\},$$

$$||F|| = \max|\nu_i| = \max\frac{1}{1-\tau\sigma\lambda_i} = \frac{1}{1+\sigma\dfrac{4K\tau}{h^2}\sin^2\dfrac{\pi}{2(2N-1)}} \leqslant 1.$$

Два слагаемых в правой части неравенства (1.5) имеют разный характер. Первое является итерационным и в простейшей ситуации $\mathrm{g}_n = 0$ получаем

$$|\delta v_n^m| \leqslant ||L||^n\,|\delta v_n^0|$$

Чтобы погрешность оставалась ограниченной, необходимо выполнение условия $||L|| \leqslant 1$. Но это выполняется, если

$$\frac{\dfrac{4K\tau}{h^2}}{1+\sigma\dfrac{4K\tau}{h^2}} \leqslant 2.$$

что равносильно неравенству

$$\left(\frac{1}{2}-\sigma\right)\frac{4K\tau}{h^2} \leqslant 1. \tag{1.6}$$

При этом условии, применяя итерационно неравенство (1.5) и учитывая, что $||F|| \leqslant 1$, получаем

$$|v_n^m| \leqslant |v_n^0| + \tau \sum_{k=0}^{m-1} |\delta \mathbf{g}_n^k| \leqslant |v_n^0| + \tau m \max_{k=\overline{0,j-1}} |\delta \mathbf{g}_n^k| \leqslant |v_n^0| + T \max_{k=\overline{0,M-1}} |\delta \mathbf{g}_n^k|,$$

что и означает устойчивость разностной схемы.

В общем случае ($K = K(u)$) анализ разностной схемы на устойчивость усложняется, так как оператор Λ уже не будут линейным. Но в конечном счете все сводится к получению оценок типа условия Липшица $|f(x) - f(y)| \leqslant C|x - y|$. Причем такие оценки можно сперва получать в малом, когда K можно считать постоянным, а затем объединять оценки по всей области. В конечном счете условие устойчивости разностной схемы в нелинейном случае получается в следующем виде:

$$\left(\frac{1}{2} - \sigma\right) \frac{4K_{\max}\tau}{h^2} \leqslant 1, \tag{1.7}$$

где $K_{\max}$ — максимальное значение коэффициента теплопроводности в области $[0, l] \times [0, T]$.

Вместо значения $K_{\max}$, которое можно определить лишь, зная искомую функцию $u(x, t)$, следует использовать какую-либо оценку сверху этой величины. Отметим, что в силу послойного характера разностной схемы в условии устойчивости можно вместо $K_{\max}$ использовать на каждом m-м шаге максимальное значение K в области $t_m \leqslant t \leqslant t_{m+1}$. Это значение близко максимальному значению $K_{\max}^m$ на прямой $t = t_m$.

Учитывая это, можно использовать, например, явную схему с переменным шагом по времени, который подбирается так, что выполняется условие устойчивости

$$\left(\frac{1}{2}-\sigma\right)\frac{4K_{\max}^{m}\tau}{h^2}\leqslant 1.$$

Это обеспечивает устойчивость разностной схемы, поскольку погрешность вычислений при переходе на очередной слой не возрастает.