

ИТЕРАЦИОННЫЕ МЕТОДЫ ДЛЯ СЕТОЧНЫХ УРАВНЕНИЙ

А. А. Самарский

Посвящается академику Л. Илиеву к его шестидесятилетию

Резюме. Рассматриваются итерационные методы (и. м.) решения операторного уравнения первого рода в гильбертовом пространстве. Теория и. м. трактуется как раздел общей теории устойчивости двухслойных и трехслойных разностных схем, развитой автором. Подробно рассмотрен т. н. попеременно-треугольный метод, основанный на поочередном обращении нижней и верхней треугольных матриц. Дано его применение к разностной задаче Дирихле для эллиптического уравнения.

Область применения численных методов в настоящее время стремительно расширяется, охватывая основные разделы физики и техники. Рассматриваемые при этом математические модели в большинстве случаев описываются при помощи линейных и нелинейных уравнений в частных производных математической физики. Поскольку каждому дифференциальному уравнению можно поставить в соответствие бесчисленное множество разностных аппроксимаций (схем), а для численного решения появляющихся при этом разностных (сеточных) уравнений — ряд методов их решения, то нет ничего удивительного в том, что в настоящее время для типичных задач математической физики предложено чрезвычайно большое число алгоритмов, предназначенных для электронно-вычислительных машин (ЭВМ). Между тем при решении конкретной задачи возникает проблема выбора одной схемы и одного алгоритма на ЭВМ; эта проблема выбора может быть решена путем применения некоторого критерия предпочтения. Фактически здесь имеется два вопроса — о выборе разностной схемы и о выборе вычислительного алгоритма для ЭВМ. Естественно требовать, чтобы теория помогла сформулировать общие принципы выбора, достаточно простые и легко проверяемые в конкретных случаях. В конечном счете алгоритм должен дать решение исходной задачи с любой заданной точностью за минимальное число операций на ЭВМ (за минимальное машинное время). Это требование экономичности алгоритма является естественным. При теоретических оценках оно часто заменяется требованием минимума арифметических действий $Q(\varepsilon)$, достаточных для получения решения с заданной точностью $\varepsilon > 0$.

Любой сеточный метод для дифференциальных и интегральных уравнений приводит к решению больших систем линейных алгебраических

уравнений, порядок которых равен числу N узлов сетки (которое в случае многомерных задач может быть большим, $N \sim 10^4 - 10^6$). Вопрос о выборе схем нужного качества на основе общей теории разностных схем достаточно подробно исследован в работах автора [1]—[2], и мы на нем не будем здесь останавливаться.

1. Рассмотрим лишь итерационные методы решения линейных сеточных уравнений. Чтобы установить иерархию методов, надо сравнить их по каким-либо характеристикам. Часто используются асимптотические оценки для числа действий при стремлении порядка системы к бесконечности (шага сетки h к нулю). Однако, несмотря на огромное быстродействие современных ЭВМ, следует признать, что существует фактически ограничение на число узлов в случае многомерных задач. Так, например, для трехмерного уравнения Лапласа ($p=3$) среднее число по каждому переменному ≈ 100 приводит к системе уравнений $N \approx 10^6$. Вряд ли целесообразно увеличение числа узлов. Поэтому надо сравнивать методы, прежде всего, на реальных (допустимых или обычно применяемых) сетках.

Часто методы рассматриваются в предположении, что процесс вычислений является идеальным (вычисления ведутся с бесконечным числом знаков). Реальный вычислительный процесс должен учитывать конечную разрядность чисел и существование ошибок округления, „машинного нуля“ и „машинной бесконечности“, с наличием которых связана возможность машинного авоста (аварийного останова). Естественно требовать, чтобы реальный вычислительный процесс был безавстным.

Чтобы выяснить качества метода на реальных сетках (для систем допустимого порядка), в том числе вопрос о безавстности, необходимо, во-первых, получить точные оценки для числа арифметических действий $Q(\varepsilon)$, во-вторых, исследовать вычислительную устойчивость метода. При этом представляет большой практический интерес характер зависимости $Q(\varepsilon)$ от точности задания априорной информации об исходной системе уравнений.

Теория итерационных методов может быть изложена как один из разделов общей теории устойчивости разностных схем, развитой в работах [1]—[3].

Характерными чертами этой теории (см. [4]—[8]) являются:

- 1) трактовка итерационных методов как операторно-разностных схем с операторами, заданными в гильбертовом пространстве;
- 2) отказ от изучения структуры операторов схемы — теория использует минимум информации общего характера относительно операторов;
- 3) конструктивный характер теории, имеющей, в конечном счете, целью указание общих принципов построения оптимальных итерационных методов в зависимости от объема априорной информации.

2. Поскольку в большинстве итерационных методов конкретная структура матрицы системы не используется, то теорию можно строить с единой точки зрения, рассматривая в качестве исходного объекта исследования операторное уравнение первого рода

$$(1) \quad Au = f.$$

Здесь A — линейный оператор, заданный на линейном конечномерном пространстве H со скалярным произведением, так что $A: H \rightarrow H$, а $f \in H$.

Относительно оператора A могут быть заданы несколько типов информации:

I. Оператор A положителен и самосопряжен,

$$(2) \quad A = A^* > 0.$$

II. $A = A^*$ и заданы границы оператора A : $\delta E \leq A \leq \Delta E$ или

$$(3) \quad \delta \|y\|^2 \leq (Ay, y) \leq \Delta \|y\|^2 \quad \forall y \in H,$$

где E — единичный оператор, $\delta > 0$ и $\Delta > 0$.

III. $A = A^* > 0$ и известны постоянные γ_1 и γ_2 энергетической эквивалентности оператора A и некоторого оператора

$$(4) \quad B = B^* > 0,$$

так что

$$(5) \quad \gamma_1 B \leq A \leq \gamma_2 B, \quad \gamma_1 > 0,$$

или $\gamma_1 (By, y) \leq (Ay, y) \leq \gamma_2 (By, y)$ для всех $y \in H$.

IV. Если A несамосопряженный оператор, то берется один из типов условий.

$$(6) \quad \text{а) } A > 0;$$

$$\text{б) } A \geq \delta E, \quad \delta > 0 \quad \|Ay\|^2 \leq \Delta (Ay, y)$$

или

$$(7) \quad A^{-1} \geq E/\Delta, \quad \Delta > 0;$$

в) заданы положительные постоянные $\gamma_1, \gamma_2, \gamma_3$, такие, что выполнение условия

$$(8) \quad \gamma_1 B \leq A_0 \leq \gamma_2 B, \quad A_0 = (A + A^*)/2$$

$$(9) \quad (B^{-1}A_1 y, A_1 y) \leq \gamma_3^2 (By, y) \quad \text{для всех } y \in H,$$

где $A_1 = (A - A^*)/2$, а $B = B^* > 0$.

Замечание 1. Из энергетических неравенств

$$(10) \quad c_1 R \leq A \leq c_2 R, \quad c_1 > 0,$$

$$(11) \quad \overset{\circ}{\gamma}_1 B \leq R \leq \overset{\circ}{\gamma}_2 B, \quad \overset{\circ}{\gamma}_1 > 0,$$

где A, B, R — положительные самосопряженные операторы, следуют неравенства (8), в которых

$$\gamma_1 = c_1 \overset{\circ}{\gamma}_1, \quad \gamma_2 = c_2 \overset{\circ}{\gamma}_2.$$

Замечание 2. Если A не является самосопряженным и знакоопределенным оператором и однородное уравнение $Au = 0$ имеет нетривиальное решение, то рассматривается обобщенное решение, т. е. решение уравнения $A'u = A^*Au = A^*f = F$, где $A' = (A')^*$.

3. Любой одношаговый итерационный метод для уравнения (1), связывающий две итерации, можно записать в каноническом виде

$$(12) \quad B_k(y_{k+1} - y_k)/\tau_{k+1} + Ay_k = f, \quad k = 0, 1, \dots, y_0 \in H,$$

где y_k — итерация (последовательное приближение) номера k , $B_k = B_k^* > 0$ — произвольный линейный оператор, заданный в H и $\tau_{k+1} > 0$, $k=0, 1, \dots$ — итерационные параметры.

Так как итерационный метод по форме совершенно аналогичен двухслойной разностной схеме (см. [2]), соответствующей абстрактной задаче Коши

$$\mathfrak{B} \frac{du}{dt} + Au = f,$$

то (12) мы будем называть двухслойной итерационной схемой (методом).

Задача ставится так: оператор A фиксирован, а параметры τ_{k+1} и оператор B_k нужно выбрать из условия минимума арифметических действий для получения решения задачи (1) с заданной точностью $\varepsilon > 0$.

Из уравнения (12) следует

$$(13) \quad y_{k+1} = y_k - \tau_{k+1} \omega_k,$$

где

$$\omega_k = B_k^{-1} r_k \text{ — поправка,}$$

$$r_k = Ay_k - f = A(y_k - u) \text{ — невязка.}$$

Применяя к (13) оператор A , получаем для поправки однородное уравнение

$$(14) \quad B_k(\omega_{k+1} - \omega_k)/\tau_{k+1} + A\omega_k = 0.$$

Таким образом, схему (12) можно трактовать как метод поправок (13), где поправка определяется из однородного уравнения (14).

В качестве условия окончания итераций можно ставить условие малости поправки:

$$(15) \quad \|\omega_n\|_D \leq \varepsilon \|\omega_0\|_D \quad (\|\omega\|_D = \sqrt{(D\omega, \omega)}),$$

где $D = D^* > 0$ — некоторый оператор; или условие

$$(16) \quad \|y_n - u\|_D \leq \varepsilon \|y_0 - u\|_D, \quad D = D^* > 0,$$

u — точное решение уравнения (1).

Недостаточно убедиться в сходимости итераций, требуется еще найти оценку для скорости сходимости итераций, т. е. для числа итераций, при котором выполнено условие (15).

Из (15) видно, что

$$(17) \quad \omega_{k+1} = S_{k+1} \omega_k, \quad S_{k+1} = E - \tau_{k+1} B_k^{-1} A$$

и, следовательно,

$$(18) \quad \omega_n = T_n \omega_0, \quad T_n = S_n S_{n-1} \dots S_1,$$

$$\|\omega_n\|_D \leq \|T_n\| \|\omega_0\|_D.$$

Условие окончания итераций означает, что норма разрешающего оператора T_n меньше ε : $\|T_n\| = q_n < \varepsilon$.

Необходимо выбрать параметры $\{\tau_k\}$ и оператор B_k так, чтобы условие $\|T_n\| = q_n < \varepsilon$ выполнялось при минимальном n , т. е. число итераций было минимальным.

В этом и состоит основная задача теории.

В частном случае, когда $A = \sum_{\alpha=1}^p A_\alpha$ есть сумма самосопряженных, положительных и попарно перестановочных операторов, применяются т. н. методы переменных направлений (МПН), для которых B_k есть факторизованный оператор (ф. о.) вида $B_k = \prod_{\alpha=1}^p (E + \omega_k^{(\alpha)} A_\alpha)$, где $\omega_k^{(\alpha)}$ — итерационные параметры, подлежащие определению. Если $p=2$, а $\delta_\alpha E \leq A_\alpha \leq \Delta_\alpha E$, $\alpha=1, 2$, то известен оптимальный набор параметров $\omega_k^{(1)}$ и $\omega_k^{(2)} = \tau_{k+1} - \omega_k^{(1)}$, при котором для получения точности ε достаточно $O(\ln 1/\eta \cdot \ln 1/\varepsilon)$ итерации, точнее $n(\varepsilon) \approx \pi^{-2} \ln 4/\eta \cdot \ln 4/\varepsilon$, где $\eta = \min_{\alpha=1, 2} \eta_\alpha$, $\eta_\alpha = \delta_\alpha/\Delta_\alpha$. Если $p > 2$, то имеются т. н. циклические итерационные параметры, при которых $n(\varepsilon) = O(\ln 1/\eta \ln 1/\varepsilon)$.

4. Для основной группы методов оператор B_k не зависит от k :

$$B_k = B = B^* > 0,$$

так что двухслойная итерационная схема имеет вид

$$(19) \quad B(y_{k+1} - y_k)/\tau_{k+1} + Ay_k = f, \quad k=0, 1, \dots, \quad \forall y_0 \in H.$$

Пусть выполнены условия II. Тогда схема (19) может быть заменена эквивалентной явной схемой

$$(20) \quad (x_{k+1} - x_k)/\tau_{k+1} + Cx_k = \varphi, \quad k=0, 1, \dots, \quad \forall x_0 \in H,$$

где $x_k = B^{1/2}y_k$, $\varphi = B^{-1/2}f$, $C = B^{-1/2}AB^{-1/2}$.

Отсюда и из (8) следует, что оператор C удовлетворяет условиям

$$(21) \quad C = C^* > 0 \quad \text{и} \quad \gamma_1 E \leq C \leq \gamma_2 E.$$

Явная схема, очевидно, соответствует уравнению

$$(22) \quad Cv = \varphi, \quad v = B^{1/2}u,$$

где u — точное решение уравнения $Au = f$. Для невязки

$$\tilde{r}_k = Cx_k - \varphi = B^{-1/2}r_k, \quad r_k = Ay_k - f,$$

получаем однородное уравнение

$$(23) \quad (\tilde{r}_{k+1} - \tilde{r}_k)/\tau_{k+1} + C\tilde{r}_k = 0 \quad \text{или} \quad \tilde{r}_{k+1} = S_{k+1}\tilde{r}_k,$$

где $S_{k+1} = E - \tau_{k+1}C$ — оператор перехода от k -й к $(k+1)$ -й итерации.

Для \tilde{r}_n находим

$$(24) \quad \tilde{r}_n = T_n \tilde{r}_0,$$

$T_n = S_n S_{n-1} \dots S_0$ — разрешающий оператор. Отсюда следует оценка

$$\|\tilde{r}_n\| \leq \|T_n\| \|\tilde{r}_0\|,$$

которой эквивалентна оценка

$$(25) \quad \|Ay_n - f\|_{B^{-1}} \leq q_n \|Ay_0 - f\|_{B^{-1}}, \quad q_n = \|T_n\|$$

для решения уравнения (14).

Разрешающий оператор $T_n(C)$ представляет собой полином степени n относительно оператора C с коэффициентами, зависящими от параметров $\tau_1, \tau_2, \dots, \tau_n$. Задача состоит в нахождении $\min_{\{\tau_k\}} \|T_n(C)\| = \min_{t \in [\gamma_1, \gamma_2]} |T_n(t)|$.

Решение этой задачи известно — $T_n(t)$ должен быть полином Чебышева. Это имеет место, если параметры τ_k выражаются через нули s_i полинома Чебышева:

$$s_i = \cos(2i-1)\pi/2n, \quad i=1, 2, \dots, n.$$

При этом коэффициент сжатия оператора T_n равен (см. [2])

$$(26) \quad q_n = 2\rho_1^n / (1 + \rho_1^n), \quad \text{где } \rho_1 = (1 - \sqrt{\xi}) / (1 + \sqrt{\xi}), \quad \xi = \gamma_1 / \gamma_2,$$

параметры τ_k определяются по формуле

$$(27) \quad \tau_k = \tau_0 / (1 + \rho_0^{\mu_k}), \quad \tau_0 = 2 / (\gamma_1 + \gamma_2), \quad \rho_0 = (1 - \xi) / (1 + \xi),$$

а μ_k — один из нулей s_i , т. е.

$$\mu_k \in \mathfrak{M}_n = \{\cos(2i-1)\pi/2n, i=1, 2, \dots, n\}.$$

Здесь \mathfrak{M}_n — множество нулей полинома Чебышева n -го порядка, упорядоченное произвольным образом.

Если положить $n=1$, т. е. взять постоянный параметр $\tau_k = \tau$, то оптимальное значение $\tau = \tau_0$, а

$$q_n = \rho_0^n$$

(схема простой итерации).

На практике (см. [9, 10, 12]) была выявлена численная неустойчивость (авоистность) метода при выборе $\mu_k = s_{k-n+1}$ или $\mu_k = s_k$, $k=1, 2, \dots, n$. Авоистность метода связана с ростом промежуточных значений y_k при $k < n$ при указанных способах упорядочения множества \mathfrak{M}_n . В работах [2, 9] для случая $n=2^p$, а затем в [11]—[12] и для произвольного n указаны способы упорядочения множества \mathfrak{M}_n , при котором чебышевский метод (19), (27) становится численно устойчивым и безавоистным.

Для числа итераций $n(\varepsilon)$ имеем $n(\varepsilon) \geq \ln \frac{2}{\varepsilon} / \ln \frac{1}{\rho}$, откуда следует

$$(28) \quad n(\varepsilon) \approx \frac{1}{2\sqrt{\xi}} \ln \frac{2}{\varepsilon}.$$

5. Для явной схемы (20) оптимизация достигается путем выбора параметров $\{\tau_k\}$. Полученный набор параметров, очевидно, пригоден и для неявной схемы (19). Остается выбрать оператор B (стабилизатор), исходя из требований экономичности и максимума отношения γ_1/γ_2 .

Для построения оператора B выбирается оператор R (регуляризатор), такой, что $R = R^* > 0$,

$$(10) \quad c_1 R \leq A \leq c_2 R, \quad c_1 > 0.$$

Оператор B строится при помощи регуляризатора R . Укажем несколько способов выбора стабилизатора B :

1) $B = R$;

2) если R представим в виде суммы p самосопряженных попарно перестановочных неотрицательных операторов

$$R = \sum_{\alpha=1}^p R_{\alpha}, \quad R_{\alpha} = R_{\alpha}^* \geq 0,$$

то в качестве B можно выбрать факторизованный оператор

$$(29) \quad B = \prod_{\alpha=1}^p (E + \omega_{\alpha} R_{\alpha});$$

3) в случае произвольного оператора R всегда можно взять в качестве B факторизованный оператор (ф. о.)

$$(30) \quad B = (E + \omega R_1)(E + \omega R_2),$$

где $R_1 = R_2^*$, $R_1 + R_2 = R$.

При таком способе факторизации операторы $B_1 = E + \omega R_1$ и $B_2 = E + \omega R_2$ имеют треугольные матрицы и являются экономичными. Неявную схему (19) с ф. о. (30) мы называем попеременно-треугольным методом (ПТМ).

Попеременно-треугольный метод с $\omega = \tau/2$ был предложен впервые автором в [8], усовершенствован в [2, 4, 6, 7] и неоднократно рассматривался в более поздних работах других авторов, например, в [15].

Ограничимся изучением ф. о. (30). Найдем постоянные γ_1, γ_2 эквивалентности операторов B и R :

$$\gamma_1 B \leq R \leq \gamma_2 B,$$

предполагая, что выполнены условия

$$R \geq \delta E, \quad R_1 R_2 \leq \Delta R/4,$$

где $\delta > 0, \Delta > 0, E$ — единичный оператор.

Представим B в виде

$$B = (E - \omega R_1)(E - \omega R_2) + 2\omega(R_1 + R_2) \geq 2\omega R,$$

так как $(E - \omega R_1)(E - \omega R_2) = (E - \omega R_1)(E - \omega R_1)^* \geq 0$. Далее имеем

$$B = E + \omega R + \omega^2 R_1 R_2/4 \leq (1/\delta + \omega + \omega^2 \Delta/4)R.$$

Из этих двух неравенств следует

$$\gamma_1 = \delta / (1 + \omega\delta + \omega^2 \Delta/4).$$

Определим ω из условия максимума функции

$$\xi(\omega) = \gamma_1 / \gamma_2 = 2\omega\delta / (1 + \omega\delta + \omega^2 \Delta/4).$$

Дифференцируя по ω , найдем точку максимума $\xi(\omega)$:

$$(31) \quad \omega = \omega_0 = 2/\sqrt{\delta\Delta}.$$

Этому значению ω соответствуют параметры

$$(32) \quad \gamma_1 = \delta/2(1 + \sqrt{\eta}), \quad \gamma_2 = \delta/4\sqrt{\eta}, \quad \xi = 2\sqrt{\eta}/(1 + \sqrt{\eta}), \quad \eta = \delta/\Delta.$$

Если $R = A$, т. е. $c_1 = c_2 = 1$, то $\gamma_1 = \gamma_1, \gamma_2 = \gamma_2$.

$$(33) \quad \xi = \gamma_1/\gamma_2 = 2\sqrt{\eta}/(1+\sqrt{\eta}),$$

а для числа итераций ПТМ с чебышевским набором параметров верна оценка

$$(34) \quad n(\varepsilon) \approx \ln \frac{2}{\varepsilon} / 2\sqrt{\xi} \approx 1 \ln \frac{2}{\varepsilon} / 2\sqrt{2\sqrt{\eta}}.$$

В случае простой итерации ($\tau_k = \tau_0$)

$$n(\varepsilon) \approx \ln \frac{2}{\varepsilon} / 2\xi \approx 1 \ln \frac{2}{\varepsilon} / 2\sqrt{\eta}.$$

Переход от y_k к y_{k+1} для схемы с ф. о. (30) можно осуществить с помощью алгоритма

$$(35) \quad \begin{aligned} (E + \omega R_1)\bar{w} &= -r_k, \quad r_k = Ay_k - f, \\ (E + \omega R_2)w_k &= \bar{w}, \quad y_{k+1} = y_k - \tau_{k+1}w_k \end{aligned}$$

или алгоритма

$$(36) \quad \begin{aligned} (E + \omega R_1)\bar{y} &= (E + \omega R_1)(E + \omega R_2)y_k - \tau_{k+1}r_k + \\ &+ (E + \omega R_2)y_{k+1} = \bar{y}. \end{aligned}$$

Первый алгоритм более экономичен, но требует запоминания не одного, а двух векторов.

6. Поясним ПТМ на модельной задаче для уравнения Пуассона. В p -мерном кубе $\bar{G} = \{0 \leq x_\alpha \leq 1, \alpha = 1, 2, \dots, p\}$ с границей Γ рассматривается задача Дирихле

$$Lu = \sum_{\alpha=1}^p L_\alpha u = -f(x), \quad x \in \bar{G}, \quad u|_\gamma = \mu(x), \quad L_\alpha u = \frac{\partial^2 u}{\partial x_\alpha^2}.$$

Пусть $\bar{\omega}_h = \{x_i = (i_1 h, \dots, i_p h) \in \bar{G}, h = 1/N\}$ — сетка с шагом $h_1 = h_2 = \dots = h_p = h$, $\gamma = \{x_i \in \Gamma\}$ — граница сетки $\bar{\omega}_h$. Формулируем разностную задачу

$$(37) \quad \begin{aligned} \Delta y &= \sum_{\alpha=1}^p \Delta_\alpha y = -f(x), \quad x \in \bar{\omega}_h, \quad y|_\gamma = \mu(x), \\ \Delta_\alpha y &= (y_{i_\alpha-1} - 2y_{i_\alpha} + y_{i_\alpha+1})/h^2. \end{aligned}$$

Обозначим $\Delta y = -\Delta y$ — оператор в пространстве H сеточных функций, заданных на $\bar{\omega}_h$ и равных нулю на границе γ_h . Операторы R_1 и R_2 определим по формулам

$$(38) \quad R_1 y = \sum_{\alpha=1}^p y_{x_\alpha}/h, \quad R_2 y = -\sum_{\alpha=1}^p y_{x_\alpha}/h,$$

причем $R = R_1 + R_2 = A$. Скалярное произведение (\cdot, \cdot) в H определяется таким образом:

$$(y, v) = \sum_{x_i \in \bar{\omega}_h} y(x_i)v(x_i)h^p.$$

Оператор $A = A^* \geq \delta E$, где,

$$\delta = \frac{4p}{h^2} \sin^2 \frac{\pi h}{2}.$$

Вычисления дают (см. [2])

$$(R_1 R_2 y, y) = \|R_2 y\|^2 = \left\| \sum_{\alpha=1}^p \frac{y_{x_\alpha}}{h} \right\|^2 \leq p \sum_{\alpha=1}^p \frac{(y_{x_\alpha}^-, y_{x_\alpha}^-)}{h^2} = \frac{p}{h^2} (Ry, y),$$

т. е.

$$\Delta = \frac{4p}{h^2}, \quad \eta = \sin^2 \frac{\pi h}{2}.$$

Если $h \ll 1$, то отсюда и из (34) следует

$$(39) \quad \xi \approx \pi h, \quad n(\varepsilon) \approx \frac{1}{2\sqrt{\pi h}} \ln \frac{2}{\varepsilon} \approx \frac{0,28}{\sqrt{h}} \ln \frac{2}{\varepsilon}$$

Число итераций для ПТМ не зависит от числа измерений p . От p зависит лишь число операций, производимых при вычислении y_{k+1} , если задана k -я итерация. Нетрудно заметить, что

$$B_1 y = (E + \omega R_1) y = \left(1 + \frac{2\omega}{h^2}\right) y - \frac{\omega}{h^2} \sum_{\alpha=1}^p y_{i_{\alpha-1}},$$

$$B_2 y = (E + \omega R_2) y = \left(1 + \frac{2\omega}{h^2}\right) y - \frac{\omega}{h^2} \sum_{\alpha=1}^p y_{i_{\alpha+1}}.$$

Отсюда и из (35) получим формулы для y_{k+1} :

$$\bar{w} = a \sum_{\alpha=1}^p \bar{w}_{i_{\alpha-1}} - b_k,$$

$$(40) \quad a = \omega / (h^2 + 2\omega p), \quad b = h^2 / (h^2 + 2\omega p),$$

$$w_k = a \sum_{\alpha=1}^p (w_k)_{i_{\alpha+1}} + b \bar{w}, \quad y_{k+1} = y_k - r_{k+1} w_k,$$

где $r_k = -\Delta y_k - f$ — невязка. Очевидно, что $w|_{\gamma=0} = 0$.

Для определения y_{k+1} надо затратить 7 умножений и $3(p+1)$ сложений на каждый узел сетки.

Вычисления w можно вести, например, вдоль строк, параллельных оси Ox_1 (при $i_1 = 1, 2, \dots, N-1$), начиная с узла с координатами $i_\alpha = 1, \alpha = 1, 2, \dots, p$, а вычисления w_k — вдоль тех же строк в обратном направлении, начиная счет с верхнего узла с координатами $i_\alpha = N-1, \alpha = 1, 2, \dots, p$.

Для метода переменных направлений (МПН) оператор $A_\alpha y = -y_{x_\alpha x_\alpha}$ и определение y_{k+1} сводится к последовательному решению трехточечных уравнений вдоль строк $Ox_\alpha, \alpha = 1, 2, \dots, p$, с помощью формул прогонки. При этом число действий на узел не зависит от числа узлов сетки и превосходит число действий на узел для ПТМ.

Представляет интерес сравнение МПН и ПТМ для разностной задачи Дирихле при $p \geq 2$ на реальных сетках с шагом $h \geq 1/100$. В случае $p = 2$

МПН с оптимальным набором параметров требует в 1,5—2 раза меньше итераций по сравнению с ПТМ с чебышевским набором параметров. В случае $p \geq 3$ для МПН не найден оптимальный набор параметров, а применяются циклические наборы [16]—[18] (наилучший из них см. в [14]).

ПТМ с чебышевскими параметрами требует при $h \geq 1/100$ для трехмерных (и тем более для $p > 3$) задач в 3—4 раза меньше итераций, чем используемые ныне методы [17, 18], и в 1,5—2 раза меньше по сравнению с [14]. Кроме того, цена одной итерации больше в МПН по сравнению с ПТМ.

7. При изучении вычислительной устойчивости итерационного метода требуется выяснить характер зависимости решения и числа итераций от изменения входных данных, т. е. от правой части f , операторов A и B , а также чисел γ_1, γ_2 .

Введение ошибок округления эквивалентно возмущению входных данных задачи, так что реальное решение задачи $\langle B, A, f, \tau_{k+1} \rangle$ можно рассматривать как точное решение следующей задачи

$$\langle \tilde{B}, \tilde{A}, \tilde{f}_k + \tilde{w}_{k+1}/\tau_{k+1}, \tau_{k+1} \rangle.$$

В работах [2], [4], [10]—[13] исследована устойчивость различных итерационных методов, а в [13] получены точные оценки.

Важным является вопрос о характере зависимости скорости сходимости от неточности задания γ_1, γ_2 . Показано, что двухслойная схема слабее зависит от возмущения γ_1 , чем трехслойная (см. [13]).

8. До сих пор предполагалось, что постоянные γ_1 и γ_2 известны. Если это не так, то можно воспользоваться вариационно-итерационными методами скорейшего спуска и минимальных невязок, не требующими знания γ_1 и γ_2 .

Изменение алгоритма в этом случае сводится к изменению формулы для параметра τ_{k+1} :

(41) а) $\tau_{k+1} = (w_k, r_k) / (Aw_k, w_k)$, $w_k = B^{-1}r_k$ — для метода скорейшего спуска;

(42) б) $\tau_{k+1} = (Aw_k, w_k) / (B^{-1}Aw_k, Aw_k)$ — для метода минимальных невязок, точнее, поправок.

Метод минимальных поправок применим, если оператор A не является самосопряженным, $A \neq A^*$; при этом справедлива оценка:

$$(43) \quad \|r_n\|_{B^{-1}} \leq \bar{q}^n \|r_0\|_{B^{-1}},$$

где

$$\bar{q} = (q_0 + \kappa) / (1 + \kappa q_0), \quad \kappa = \gamma_3 / \sqrt{\gamma_1 \gamma_2 + \gamma_3^2},$$

$$q_0 = (1 - \xi) / (1 + \xi), \quad \xi = \gamma_1 / \gamma_2,$$

если выполнены условия

$$A = A_0 + A_1, \quad A_0 = \frac{1}{2}(A + A^*), \quad A_1 = \frac{1}{2}(A - A^*),$$

$$\gamma_1 B \leq A_0 \leq \gamma_2 B, \quad (B^{-1}A_1 y, A_1 y) \leq \gamma_3^2 (B y, y)$$

или $A_1^* B^{-1} A_1 \leq \gamma_3^2 B$, $\gamma_1 > 0$, $\gamma_3 > 0$.

Если $A=A^*>0$, то $\bar{q}=q_0$ и $q_n=q_0^n$, так что $n(\varepsilon)\approx\frac{1}{2\xi}\ln\frac{1}{\varepsilon}$ как для метода скорейшего спуска, так и для метода минимальных поправок.

9. Рассмотрим случай, когда $\lambda=0$ является собственным значением оператора $A=A^*\geq 0$, которому соответствуют ортонормированные собственные векторы $\xi_1, \xi_2, \dots, \xi_m$. Пусть $H^{(0)}$ — линейное подпространство размерности m с базисом ξ_1, \dots, ξ_m , $H^{(1)}$ — дополнение $H^{(0)}$ до H . Любой вектор $y \in H^{(0)}$ удовлетворяет уравнению $Ay=0$. Можно указать несколько способов решения задачи $Ay=f$ в этом случае. Оператор $B=B^*>0$ отображает H в H . Рассмотрим схему (19).

Первый способ. Если $f \perp H^{(0)}$, т. е. $f \in H^{(1)}$, то все итерации $y_k \in H^{(1)}$, при условии, что нулевое приближение $y_0 \in H^{(1)}$. Справедливы неравенства

$$\gamma_1 \|x\|_B^2 \leq (Ax, x) \leq \gamma_2 \|x\|_B^2 \quad \forall x \in H^{(1)}, \quad \gamma_1 > 0.$$

В качестве $\{\tau_k\}$ можно взять устойчивый набор чебышевских параметров, используя постоянные $\gamma_1 > 0$ и $\gamma_2 > 0$. Из-за ошибок округления условие $f \in H^{(1)}$ может нарушаться и потому нуждается в контроле.

Второй способ. Представим f в виде $f=f^{(0)}+f^{(1)}$, где $f^{(0)} \in H^{(0)}$, $f^{(1)} \in H^{(1)}$, и предположим, что $f^{(0)} \neq 0$, т. е. не выполнено условие разрешимости. Тогда решение является обобщенным и определяется из уравнения

$$A^2u = Af = F,$$

причем условие согласования выполняется автоматически, так как $F \in H^{(1)}$. Будем пользоваться итерационной схемой с параметрами, выбираемыми для A как оператора в $H^{(1)}$.

Для поправки $w_k = B^{-1}r_k$ имеем уравнение

$$B(w_{k+1} - w_k)/\tau_{k+1} + Aw_k = 0.$$

Рассмотрим для простоты случай простой итерации, когда $\tau_{k+1} = \tau_0 = 2/(\gamma_1 + \gamma_2)$, где $\gamma_1 > 0$. Представим y_k и f в виде $y_k = y_k^{(0)} + y_k^{(1)}$, $f = f^{(0)} + f^{(1)}$, $f^{(i)} \in H^{(i)}$, $i=0,1$. Из уравнения (19) следует, что

$$B(y_{k+1}^{(0)} - y_k^{(0)})/\tau_0 = f^{(0)}, \quad y_{k+1}^{(0)} = y_k^{(0)} - \tau_0 B^{-1}f^{(0)}.$$

Суммирование по $k=0, 1, 2, \dots$ дает

$$y_k^{(0)} = y_0^{(0)} - \tau_0 B^{-1}f^{(0)}k.$$

Таким образом,

$$y_k = y_k^{(1)} + y_k^{(0)} = y_k^{(1)} + y_0^{(0)} - \tau_0 B^{-1}f^{(0)}k.$$

Отсюда следует, что вектор

$$v_k = y_k - k(y_{k+1} - y_k) = y_k^{(1)} + y_0^{(0)} - k(y_{k+1}^{(1)} - y_k^{(1)})$$

принадлежит $H^{(1)}$, если $y_0^{(0)} = 0$ (И. Н. Молчанов и М. Яковлев).

Если u — решение задачи $A^2u = Af$, то разность $v_k - u$ также сходится к нулю при $k \rightarrow \infty$, однако скорость сходимости несколько меньше, чем для $y_k^{(1)} - u$ (вместо $q_k = e_0^k$ получаем $\tilde{q}_k = e_0^k(k-1 + k e_0) = q_k(k(1+e_0)-1)$).

Указанный выше метод „сбрасывания“ компонент из $H^{(0)}$ может быть распространен на случай чебышевского набора параметров $\{\tau_k\}$:

$$u_k = y_k - \alpha_k (y_{k+1} - y_k), \quad \alpha_k = \sum_{i=1}^{k+1} \tau_i / \tau_{k+1}.$$

Впрочем, можно не производить регуляризацию решения переходом от y_k к u_k . Так как $B > 0$, то задача (19) всегда имеет единственное решение, содержащее компоненту $y_{k+1}^{(0)}$. Если она не слишком велика, то можно в качестве приближенного решения задачи $Au = f$ взять найденную итерацию y_{k+1} .

ЛИТЕРАТУРА

1. А. А. Самарский. О регуляризации разностных схем. *Ж. вычисл. мат. и мат. физ.*, 7, 1967, № 1, 62—93.
2. А. А. Самарский. Введение в теорию разностных схем. Москва, 1971.
3. А. А. Самарский. Классы устойчивых схем. *Ж. вычисл. мат. и мат. физ.*, 7, 1967, № 5, 1096—1133.
4. А. А. Самарский. Двухслойные итерационные схемы. *Доклады АН СССР*, 185, 1969, № 3, 524—527.
5. А. А. Самарский. Итерационные двухслойные схемы для несамосопряженных уравнений. *Доклады АН СССР*, 186, 1969, № 1, 35—38.
6. А. А. Самарский. Некоторые вопросы теории разностных схем. *Ж. вычисл. мат. и мат. физ.*, 6, 1966, № 4, 665—686.
7. А. А. Самарский. Некоторые вопросы общей теории разностных схем. В сб.: Дифференциальные уравнения с частными производными. Москва, 1970, 191—223.
8. А. А. Самарский. Об одном экономичном алгоритме численного решения систем дифференциальных и алгебраических уравнений. *Ж. вычисл. мат. и мат. физ.*, 4, 1964, № 3, 580—585.
9. В. Вазон, Дж. Форсайт. Разностные методы решения дифференциальных уравнений с частными производными. Москва, 1963.
10. В. И. Лебедев, С. А. Финогенов. О порядке выбора итерационных параметров в циклическом процессе. *Ж. вычисл. мат. и мат. физ.*, 12, 1972, № 2, 427—438.
11. Э. М. Михайлов, А. А. Самарский. Выбор итерационных параметров в методе Рундсона. *Ж. вычисл. мат. и мат. физ.*, 12, 1972, № 4, 960—973.
12. В. И. Лебедев, С. А. Финогенов. Решение проблемы упорядочения параметров в чебышевских итерационных методах. *Ж. вычисл. мат. и мат. физ.*, 13, 1973, № 1, 18—33.
13. Е. С. Николаев, А. А. Самарский. О вычислительной устойчивости двухслойных и трехслойных итерационных схем. *Ж. вычисл. мат. и мат. физ.*, 12, 1972, № 5, 1197—1207.
14. Е. С. Николаев, А. А. Самарский. Методы численного решения задачи Дирихле для уравнения Пуассона любого числа измерений. *Доклады АН СССР*, 206, 1972, № 4, 815—818.
15. В. П. Ильин. О некоторых численных экспериментах по итерационным методам решения разностных уравнений Лапласа. Численные методы механики сплошной среды (информационный бюллетень, Новосибирск), 1:6 (1970), 31—51.
16. J. Douglas. Alternating direction method for three space variables. *Numer. Math.*, 4, 1962, No 1, 41—63.
17. А. А. Самарский, В. Б. Андреев. Итерационные схемы переменных направлений для численного решения задачи Дирихле. *Ж. вычисл. мат. и мат. физ.*, 4, 1964, № 6, 1025—1036.
18. A. Nadjidimos. Optimum extrapolated ADI iterative difference schemes for the solution of Laplace's equation in three space variables. *Comput. J.*, 14, 1971, No 2, 179—183.