Записки научных
семинаров ПОМИ
Том 496, 2020 г.

В. П. Ильин

ПАРАЛЛЕЛЬНЫЕ ПОСМЕННО ТРЕУГОЛЬНЫЕ ИТЕРАЦИОННЫЕ МЕТОДЫ В ПОДПРОСТРАНСТВАХ КРЫЛОВА

§1. Введение

Целью данной работы является построение и исследование параллельных итерационных методов в подпространствах Крылова для решения систем линейных алгебраических уравнений (СЛАУ)

$$Au = f, \quad A \in \mathcal{R}^{N,N}, \quad u, f \in \mathcal{R}^{N}, \tag{1}$$

с вещественными разреженными матрицами высокого порядка ($N \approx 10^{10}$ и выше), имеющими большие числа обусловленности (10^{13} и более), реализация которых на современных многопроцессорных вычислительных системах (МВС) представляет актуальную практическую проблему. В частности, при решении прямых и обратных междисциплинарных задач математического моделирования с реальными данными, в том числе нелинейных и нестационарных, данная расчетная стадия может занимать около 80 % времени машинного эксперимента, поскольку здесь объемы затрачиваемых компьютерных ресурсов растут нелинейно с увеличением числа степеней свободы.

Нас интересуют главным образом СЛАУ, которые возникают при аппроксимации многомерных начально-краевых задач, характеризуемых переменными коэффициентами и контрастными материальными свойствами, с помощью методов конечных разностей, конечных объёмов, конечных элементов или разрывных алгоритмов Галеркина [1]. Предполагается, что в таких случаях непосредственно неприменимы специальные рекордные методы типа быстрого преобразования Фурье. Основные подходы при этом базируются на предобусловленных итерационных алгоритмах в подпространствах Крылова. Типичная форма

Ключевые слова: посменно треугольные матрицы, вложенные факторизации, методы сопряженных направлений, распараллеливание вычислений.

Работа выполнена при финансовой поддержке РФФИ (грант No. 18-01-00295.)

легко обращаемой предобусловленной матрицы заключается в приближенной факторизации вида

$$B = (G + L)G^{-1}(G + U) = G + L + U + LG^{-1}U, \tag{2}$$

где L и U – нижняя и верхняя блочно треугольные части исходной матрицы $A = D + L + U$, а G, D – некоторые невырожденные блочно диагональные матрицы. Отметим, что если A – блочно трехдиагональная матрица и матрица G в (2) определяется из матричного уравнения

$$G = D - LG^{-1}U, \tag{3}$$

то данное представление является точной блочной факторизацией матрицы A.

На основе приближенных аналогов формулы (3) конструируются различные методы симметричной последовательной верхней релаксации (SSOR), а также явной или неявной неполной факторизации (в частности, алгоритмы ILU – неполного треугольного разложения), см. обзоры в [2–7].

Рассматриваемые методы эффективно применяются для решения блочно трехдиагональных СЛАУ, актуальных во многих приложениях. Например, в двумерной прямоугольной области на прямоугольной сетке (возможно, неравномерной) с количеством расчетных узлов $N = N_x N_y$ (N_x и N_y – числа шагов по координатам x, y соответственно) при их естественной упорядоченности стандартные пяти- или девятиточечные аппроксимации уравнений диффузионного типа [1] порождают алгебраические системы вида

$$Au = \begin{bmatrix} D_1 & U_1 & & & 0 \\ L_2 & D_2 & U_2 & & \\ & \ddots & \ddots & \ddots & \\ & & & & U_{N_x-1} \\ 0 & & & L_{N_x} & D_{N_x} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \cdot \\ \\ u_{N_x} \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ \cdot \\ \\ f_{N_x} \end{bmatrix}. \tag{4}$$

Матрицы $D = \{D_i\}, L = \{L_i\}, U = \{U_i\}$ здесь являются блочно диагональными ($D_i, L_i \in \mathcal{R}^{N_y,N_y}$), а матрица A предполагается симметричной положительно определенной (с.п.о). Возможные обобщения будут рассмотрены ниже отдельно. Заметим, что для блочно трехдиагональной структуры (4) матрицы A ее точная факторизация представляется формулами (2), (3), сводящимися в данном случае к блочным прогонкам и содержащим блочно диагональную матрицу $G = \text{block-diag}\{G_i\}$, в которой блоки G_i определяются рекурсивным образом. Поскольку реализация такого алгоритма требует трудоёмкого обращения матриц G_i, более предпочтительным является построение экономичных итерационных методов, использующих приближенную ленточную аппроксимацию обратных к ленточным матрицам. При этом улучшение предобуславливания СЛАУ удаётся осуществить с помощью обобщенного согласования строчных сумм исходной и предобуславливающей матриц, см. [2, 5, 8]:

$$G_1 = D_1, \quad \bar{G}_i = D_i - L_i \bar{G}_{i-1}^{-1} U_{i-1} - \theta S_i, \quad i = 1, 2, ..., N_x. \tag{5}$$

Здесь $\bar{G}_{i-1}^{-1}$ означает трехдиагональную часть матрицы G_{i-1}^{-1} (её главную и соседние к ней диагонали), $\theta \in [0,1]$ – некоторый компенсирующий параметр, а S_i – диагональные матрицы, определяемые из принципа обобщенного согласования строчных сумм, или фильтрации:

$$Sy^{(l)} = [L(G^{-1} - (\bar{G}^{-1})U]y^{(l)}, \quad l = 1, m,$$

где $S = \text{block-diag}\{S_i\}, \bar{G} = \text{diag}\{\bar{G}_i^{-1} = 3\text{-diag}\{G_i^{-1}\}\}$, а $y^{(l)}$ – пробные, или фильтрующие, векторы, $m = 1$ или $m = 2$.

Более простой способ построения предобуславливателя заключается в выборе, вместо (5), $G_i = \omega^{-1} D_i$, что приводит к блочному методу симметричной последовательной верхней релаксации (BSSOR [2, 5]), оптимальное значение релаксирующего параметра ω которого ищется в полуоткрытом интервале [1, 2).

Если расчетная область трехмерная и имеет форму параллелепипеда, а сетка обладает числом узлов $N = N_x N_y N_z$ (N_z – количество шагов вдоль оси z), то соответствующая СЛАУ также может быть представлена в форме (4), но матрицы D_i, L_i, U_i при этом будут иметь размерность $N_y N_z$.

Обозначим через h характерный шаг сетки и предположим, что она является регулярной, или квазиравномерной, т. е. при $h \to 0$ все шаги сетки являются величинами одного порядка. Тогда число обусловленности матрицы A имеет порядок

$$\text{cond}(A) = \max\{\lambda(A)\}/\min\{\lambda(A)\} = O(h^{-2}),$$

причем оно ухудшается (увеличивается) при усилении неравномерности сетки и контраста, или разномасштабности, коэффициентов исходного решаемого уравнения. При использовании матрицы B вида (2) число обусловленности получаемой предобусловленной СЛАУ удается улучшить на порядок, т. е. $\text{cond}(B^{-1}A) = O(h^{-1})$. После применения итерационных алгоритмов в подпространствах Крылова решение алгебраических систем можно получить за $n = O(h^{-1/2})$ последовательных приближений.

Эта высокая эффективность ослабляется тем, что для рассматриваемых алгоритмов трудно повысить производительность на многопроцессорных вычислительных системах (МВС), поскольку обращение предобуславливающей матрицы B требует решения плохо распараллеливаемых треугольных алгебраических систем.

В работе [9] был предложен двухпотоковый блочный вариант посменно треугольной факторизации исходной матрицы, когда каждый из множителей в предобуславливателе B не является нижней или верхней треугольной матрицей, а состоит из блочных строк различной ориентации: одни являются нижними треугольными, а остальные – верхними треугольными (такое разложение названо авторами “twisted decomposition”, а в русскоязычной литературе данный подход традиционно определяется как алгоритм встречных прогонок, см. обзор в [10]).

Мы рассмотрим обобщение данного приема, определяя матрицы L и U как посменно треугольные, т.е. состоящие из P блочных строк, которые последовательно меняют свою принадлежность к нижне- или к верхнетреугольной части матрицы. При этом решение СЛАУ с матрицами $G + L$ и $G + U$ может быть распараллелено на P процессорах.

Настоящая работа построена следующим образом. В §2 описываются посменно треугольные итерационные методы блочной симметричной последовательной релаксации и неполной факторизации в подпространствах Крылова для решения сеточных двумерных краевых задач, описываемых блочно трехдиагональными СЛАУ вида (4). §3 посвящен обобщению предложенного подхода к распараллеливанию трехмерных сеточных алгебраических систем путем использования методов вложенной факторизации (nested factorization), предложенных в работах [11, 12], а затем развиваемых различными авторами [13–15] и широко применяемых в программном обеспечении специалистами нефтегазовой отрасли [16]. В §4 описываются посменно треугольные предобусловленные методы в подпространствах Крылова трех типов (алгоритмы сопряженных градиентов, сопряженных невязок и минимальных ошибок) с различными вариационными и ортогональными свойствами. В заключении мы обсуждаем перспективные направления исследований как в отношении повышения скорости сходимости итераций, так и производительности рассматриваемых параллельных алгоритмов.

§2. Посменно треугольные методы решения сеточных двумерных задач

При использовании матрицы B в качестве предобуславливателя для какого-либо итерационного процесса на каждом шаге требуется решать вспомогательную систему следующего вида (см. ниже (32) в §4):

$$Bp \equiv (G+L)G^{-1}(G+U)p = r. \tag{6}$$

Ее решение можно найти из последовательных соотношений

$$(G+L)v = r, \quad (G+U)p = Gv = w. \tag{7}$$

Как известно, решение СЛАУ вида (7) с треугольными матрицами L, U плохо распараллеливается на МВС. В качестве альтернативы мы рассмотрим матрицы и методы посменно треугольного типа. Определение посменно треугольных матриц алгоритмов решения соответствующих СЛАУ на примере блочно трехдиагональных систем вида (5) проиллюстрируем сначала для случая блочного порядка $N_x = 7$ и одной смены "треугольности":

$$(G+L)v \equiv \left[\begin{array}{cc|c|c|c|c|c} G_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ L_2 & G_2 & 0 & 0 & 0 & 0 & 0 \\ 0 & L_3 & G_3 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & L_4 & G_4 & U_4 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & G_5 & U_5 & 0 \\ 0 & 0 & 0 & 0 & 0 & G_6 & U_6 \\ 0 & 0 & 0 & 0 & 0 & 0 & G_7 \end{array}\right] \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \\ v_7 \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \\ r_5 \\ r_6 \\ r_7 \end{bmatrix} = r, \tag{8}$$

$$(G+U)p \equiv \left[\begin{array}{ccc|c|c|c|c} G_1 & U_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & G_2 & U_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & G_3 & U_3 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & G_4 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & L_5 & G_5 & 0 & 0 \\ 0 & 0 & 0 & 0 & L_6 & G_6 & 0 \\ 0 & 0 & 0 & 0 & 0 & L_7 & G_7 \end{array}\right] \begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \\ p_5 \\ p_6 \\ p_7 \end{bmatrix} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \\ w_5 \\ w_6 \\ w_7 \end{bmatrix} = Gv. \tag{9}$$

В соотношениях (8), (9) матрицы L, U определены иначе, чем в уравнении (3), но их сумма $L+U$ в обоих случаях одинакова.

При такой структуре матриц $G+L$ и $G+U$ формула (7) для G, обеспечивающая точное равенство $B=A$, остается без изменений. Однако рекуррентное последовательное соотношение (5) для вычисления блоков G_i в данном случае существенно изменяется. А именно, их нахождение реализуется встречными прогонками от краёв к центру по следующим формулам:

$$\begin{aligned}
G_1 &= D_1, \quad G_i = D_i - U_i \bar{G}_2^{-1} L_2 - \theta S_i, \\
S_i y^{(l)} &= L_i (G_{i-1}^{-1} - \bar{G}_{i-1}^{-1}) U_{i-1} y^{(l)}, \quad i = 2, \dots, m, \\
G_{N_x} &= D_{N_x}, \quad G_i = D_i - U_i \bar{G}_{i+1}^{-1} L_{i+1} - \theta S_i, \\
S_i y^{(l)} &= U_i (G_{i+1}^{-1} - \bar{G}_{i+1}^{-1}) L_{i+1} y^{(l)}, \quad i = N_x - 1, \dots, m+2, \\
G_{m+1} &= D_{m+1} - L_{m+1} \bar{G}_m^{-1} U_m - U_{m+1} \bar{G}_{m+2}^{-1} L_{m+2} - \theta S_{m+1}, \\
S_{m+1} y^{(l)} &= [L_{m+1} (G_m^{-1} - \bar{G}_m^{-1}) U_m + U_{m+1} (G_{m+2}^{-1} - \bar{G}_{m+2}^{-1}) L_{m+2}] y^{(l)}.
\end{aligned}$$

Естественно, все $2m+1$ трехдиагональные матрицы G_i вычисляются один раз до итерации и запоминаются. При этом прямая $(i = 1, ..., m)$ и обратная $(i = 2m+1, ..., m+2)$ прогонки легко распараллеливаются на двух потоках, или процессорных ядрах.

Решение каждого из уравнений (8), (9) можно осуществлять параллельно по следующим схемам. Из (8) блочные неизвестные вычисляются в таком порядке: сначала синхронно находятся $v_1 = G_1^{-1} r_1, v_i = G_i^{-1}(r_i - U_i v_{i+1})$ для $i = 2, 3$ и $v_7 = G_7^{-1} r_7, v_i = G_i^{-1}(r_i - L_i v_{i-1})$ для $i = 6, 5$, а в конце определяется $v_4 = G_4^{-1}(r_4 - L_4 r_3 - U_4 r_5)$. По аналогичным формулам, но в другом порядке, решается СЛАУ (9): сначала рассчитывается p_4, затем параллельно p_3, p_2, p_1 и p_5, p_6, p_7.

Чтобы перейти от конкретной иллюстрации к общему описанию посменно треугольных методов, надо формализовать крупноблочную структуру матриц $G+L, G+U = (G+L)^\top$ для достаточно больших значений блочного порядка N_x.

Каждую из матриц в СЛАУ вида (8), (9) можно представить как состоящую из двух диагональных блоков одинаковой (для простоты) блочной размерности m, отделенных друг от друга блочными строками и столбцами. Эти "крупные" блоки соответственно являются нижним (левым для (8) и правым в (9)) и верхним треугольными. Определим периодическую структуру как пару матричных блоков, левый верхний из которых является нижним треугольным, второй –верхним треугольным, а разделены они “крестом”, образованным трехдиагональной блочной строкой и диагональным блочным столбцом. Таким образом, блочный порядок этих матриц равен $2m+1$.

Перейдем теперь к более общему случаю многосменной триангуляции и предположим, что матрицы A, $G+L$ и $G+U$ имеют периодическую блочно трехдиагональную структуру с блочным порядком $N_x = (2m+1)P + P - 1$. Матрицы с двусменной триангуляцией вида (8), (9) будут в этом случае диагональными блоками, каждый из которых имеет порядок $2m+1$ и отделяется от соседних “крестом” из одной строки и одного столбца. Обозначая эти “главные” блоки через $\widehat{H}_s$ и $\check{H}_s$ в $G+L$ и $G+U$ соответственно, $s = 1, 2, \ldots, P$, получаемые СЛАУ мы можем записать как

$$(G{+}L)v = \begin{bmatrix} \widehat{H}_1 & \widehat{U}_1 & 0 & 0 & & & 0 \\ 0 & \bar{G}_1 & 0 & 0 & & & \\ 0 & \widehat{L}_2 & \widehat{H}_2 & \widehat{U}_2 & & & \\ 0 & 0 & 0 & \bar{G}_2 & \ddots & & \\ & & & \ddots & \ddots & \ddots & \\ & & & & \ddots & \bar{G}_{P-1} & 0 \\ 0 & & & & 0 & \widehat{L}_P & \widehat{H}_P \end{bmatrix} \begin{bmatrix} \widehat{v}_1 \\ \bar{v}_1 \\ \widehat{v}_2 \\ \bar{v}_2 \\ \vdots \\ \bar{v}_{P-1} \\ \widehat{v}_P \end{bmatrix} = \begin{bmatrix} \widehat{r}_1 \\ \bar{r}_1 \\ \widehat{r}_2 \\ \bar{r}_2 \\ \vdots \\ \bar{r}_{P-1} \\ \widehat{r}_P \end{bmatrix}, \quad (10)$$

$$(G{+}U)p = \begin{bmatrix} \check{H}_1 & 0 & 0 & 0 & & & 0 \\ \check{L}_1 & \bar{G}_1 & \check{U}_1 & 0 & & & \\ 0 & 0 & \check{H}_2 & 0 & & & \\ 0 & 0 & \check{L}_2 & \bar{G}_2 & \check{U}_2 & & \\ & & \ddots & \ddots & \ddots & & \\ & & & \check{L}_{P-1} & \bar{G}_{P-1} & \check{U}_{P-1} & \\ 0 & & & 0 & 0 & \check{H}_P & \end{bmatrix} \begin{pmatrix} \check{p}_1 \\ \bar{p}_1 \\ \check{p}_2 \\ \bar{p}_2 \\ \vdots \\ \bar{p}_{P-1} \\ \check{p}_P \end{pmatrix} = \begin{bmatrix} \check{w}_1 \\ \bar{w}_1 \\ \check{w}_2 \\ \bar{w}_2 \\ \vdots \\ \bar{w}_{P-1} \\ \check{w}_P \end{bmatrix}. \quad (11)$$

Здесь $\bar{G}_s = G_{q_s}$, $q_s = (2m{+}1)s + 1$ есть “обычная” трехдиагональная матрица из (9) порядка N_y, а матрицы $\widehat{L}_s, \check{L}_s, \widehat{U}_s$ и $\check{U}_s$ суть блочные строки и столбцы, определяемые следующим образом:

$$\widehat{L}_s = \begin{bmatrix} L_{q_{s+1}} \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \widehat{U}_s = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ U_{q_{s-1}} \end{bmatrix} \in \mathcal{R}^{2m+1, N_y}, \quad \begin{array}{l} \check{L}_s = [0 \ \ldots \ 0 \ L_{q_s}] \in \mathcal{R}^{N_y, 2m+1}, \\ \check{U}_s = [U_{q_s} \ 0 \ \ldots \ 0]. \end{array} \quad (12)$$

Векторы $\widehat{v}_s, \widehat{r}_s, \check{p}_s, \check{w}_s$, как видно в (10), (11) из контекста, суть векторы размерности $(2m+1)N_y$, а каждый из векторов $\bar{v}_s, \bar{r}_s, \bar{p}_s, \bar{w}_s$ имеет порядок N_y.

Решение алгебраических систем (10) и (11) на МВС может осуществляться параллельно, но по разным вычислительном схемам. В СЛАУ (10) сначала синхронно вычисляются посменные разделители $\bar{v}_s = \bar{G}_s^{-1}\bar{r}_s$, $s = 1, 2, \ldots, P-1$, а затем находятся крупноблочные компоненты из P систем вида

$$\widehat{H}_s \widehat{v}_s = \widehat{r}_s - \widehat{L}_s \bar{v}_{s-1} - \widehat{U}_s \bar{v}_{s+1}. \tag{13}$$

При этом каждая s-я система (13) решается аналогично СЛАУ (8), т. е. с помощью встречных прогонок от краев к центру.

Уравнения вида (11) решаются в противоположной последовательности: сначала из систем

$$\check{H}_s \check{p}_s = \check{w}_s, \quad s = 1, 2, \ldots, P, \tag{14}$$

одновременно определяются “крупные” подвекторы $\check{p}_s$. В этом случае, аналогично реализации (9), применяются встречные прогонки “от центра к краям”. На заключительном этапе компоненты-разделители находится из СЛАУ

$$\bar{G}_s \bar{p}_s = \bar{w}_s - \bar{L}_s \check{p}_{s-1} - \bar{U}_s \check{p}_{s+1}, \quad s = 1, 2, \ldots P-1. \tag{15}$$

§3. Посменно треугольные алгоритмы для трехмерных задач

При решении трехмерных краевых задач мы останавливаемся на стандартных семиточечных аппроксимациях уравнений диффузионного типа при использовании параллелепипедальных сеток. Матрица системы уравнений порядка $N = N_x N_y N_z$ вместо (4) может быть представлена в форме

$$A = D + L_1 + U_1 + L_2 + U_2 + L_3 + U_3. \tag{16}$$

Здесь D есть главная диагональ исходной матрицы, а L_l и U_l, $l = 1, 2, 3,$ – это диагонали матрицы A, относящиеся к ее нижней и верхней частям. Индексы $l = 1, 2, 3$ можно трактовать как соответствующие декартовым переменным x, y, z при трехточечных аппроксимациях вторых производных решаемого дифференциального уравнения по различным направлениям.

В данном случае, мы будем использовать методы вложенных факторизаций, которые определяются на основе построения предобуславливающей матрицы B вида (2) следующим рекурсивным образом (см. [11–16]):

$$\begin{aligned} B &= (P + L_3)P^{-1}(P + U_3) = P + L_3 + U_3 + L_3P^{-1}U_3, \\ P &= (T + L_2)T^{-1}(T + U_2) = T + L_2 + U_2 + L_2T^{-1}U_2, \\ T &= (M + L_1)M^{-1}(M + U_1) = M + L_1 + U_1 + L_1M^{-1}U_1, \end{aligned} \tag{17}$$

в результате чего получаем

$$B = M + A - D + L_1M^{-1}U_1 + L_2T^{-1}U_2 + L_3P^{-1}U_3. \tag{18}$$

При естественной упорядоченности узлов сетки и соответствующих векторных компонент матрицы M, T и P формируются как диагональная, трехдиагональная и пятидиагональная, а предобуславливатель B определяется формулами

$$\begin{aligned} M &= D - L_M^{-1}U_1 - \theta_1S_1 - \theta_2S_2, \\ B &= A + L_2T^{-1}U_2 + L_3P^{-1}U_3 - \theta_1S_1 - \theta_2S_2. \end{aligned} \tag{19}$$

Здесь θ_1 и θ_2 – итерационные (релаксирующие, или компенсирующие) параметры, а S_1 и S_2 – диагональные матрицы, определяемые из принципа согласования строчных сумм ($Ae = Be$):

$$S_1e = L_2T^{-1}U_2e, \quad S_2e = L_3P^{-1}U_3e, \tag{20}$$

где $e = \{1\}$ – вектор с единичными компонентами.

Отметим, что в (19) матричное уравнение для M однозначно разрешимо в классе диагональных матриц, т. е. M имеет ту же структуру, что и D.

В формулах (17)–(20) метода вложенных факторизаций матрицы L_l, $l = 1, 2, 3$, являются нижними треугольными, а U_l – верхними треугольными. Если же матрицы L_2, L_3, U_2 и U_3 определить, как в предыдущем параграфе, то мы получим посменно треугольный вариант алгоритма вложенных факторизаций. Заметим, что для матриц L_1 и U_1 имеется ограничение: их следует определить как односменно треугольные, причем в L_1 левый блок должен быть нижним треугольным. Только в этом случае матричное уравнение для M из (19) разрешимо в классе диагональных матриц. Фактически реализация факторизации по первому направлению сводится к скалярной версии встречных прогонок, которые можно реализовать на $P_1 = 2$ вычислительных потоках. Посменно треугольные матрицы по второму и третьему направлениям можно определять как многосменные ($P_2 \leqslant N_y, P_3 \leqslant N_z$). Таким образом, данный вариант вложенных факторзаций допускает распараллеливание на $P = 2P_2P_3$ потоках.

Если предобуславливатель B вида (17), (18) применить к какому-либо итерационному процессу в подпространствах Крылова, то мы получаем трехуровневую факторизацию в классическом одноуровневом итерационном методе. Однако при решении рассматриваемых трехмерных сеточных краевых задач можно поступить и по-другому, формируя двухуровневую факторизацию и двухуровневый итерационный процесс.

Для этого перепишем исходную матрицу A из (16) в виде

$$A = D_3 + L_3 + U_3, \quad D_3 = D_2 + L_2 + U_2, \quad D_2 = D + L_1 + U_1. \tag{21}$$

В данном случае, D_3 есть блочно диагональная матрица с пятидиагональными блоками $D_{3,i}$ размерности N_yN_z, каждый из которых соответствует плоской задаче в сечении $x = \text{const}$ и имеет структуру матрицы A в (6). Тогда матрицу P в (17) определяем формулой

$$P = \{G_i = D_{3,i} - \theta S_i\}, \quad S_ie = L_{3,i}G_{i-1}^{-1}U_{i-1}e, \tag{22}$$

что соответствует определению $\overline{L_3P^{-1}U_3} = 0$ в (17). Отметим, что L_3 и U_3 можно определить как посменно треугольные, и тогда реализацию алгоритма можно распараллеливать с помощью встречных блочных прогонок.

Если в (22) положить $\theta = 0$, то мы придем к блочному методу симметричной последовательной верхней релаксации (SSOR, [4]). В этом случае, в (17) надо заменить P на $\omega^{-1}P$, где релаксационный параметр ω имеет оптимальное значение на отрезке $[1, 2]$.

Отметим, что каждая вспомогательная двумерная СЛАУ обладает строгим диагональным преобладанием, имеет конечное число обусловленности, а границы ее спектра можно оценить с помощью кругов Гершгорина.

§4. Посменно треугольные методы в подпространствах Крылова

Рассмотрение итерационных алгоритмов мы будем проводить в применении к СЛАУ, полученной из исходной системы вида (1) путем двустороннего предобуславливания:

$$\bar{A}\bar{u} = \bar{f}, \quad \bar{A} = L_B^{-1} A U_B^{-1}, \quad \bar{u} = U_B u, \quad \bar{f} = L_B^{-1} f. \tag{23}$$

Здесь L_B и U_B – множители из факторизации невырожденной предобуславливающей матрицы

$$B = L_B U_B, \quad B^{-1} = U_B^{-1} L_B^{-1}. \tag{24}$$

Для решения предобусловленной СЛАУ $\bar{A}\bar{u} = \bar{f}$ с симметричной положительно определенной матрицей $\bar{A} = \bar{A}^\top$ рассмотрим итерационные процессы следующего вида:

$$\begin{gathered} \bar{u}^{n+1} = \bar{u}^n + \alpha_n p^n = \bar{u}^0 + \alpha_0 p^n + \cdots + \alpha_n p^n, \\ \bar{r}^{n+1} = \bar{f} - \bar{A}\bar{u}^{n+1} = \bar{r}^n - \alpha_n \bar{A} p^n = \bar{r}^0 - \alpha_0 \bar{A} p^0 - \cdots - \alpha_n \bar{A} p^n, \end{gathered} \tag{25}$$

где p^n – некоторые направляющие векторы, α_n – итерационные параметры, $\bar{u}^0 = U_B u^0$ и $\bar{r}^0 = \bar{f} - \bar{A}\bar{u}^0$ – предобусловленные векторы начального приближения и невязки, а u^0 – произвольный вектор.

В формулах (25) предполагается, что векторы p^n обладают следующими свойствами ортогональности:

$$\left(\bar{A}^\gamma p^n, p^k\right) = (p^n, p^k)_\gamma = \rho_n^{(\gamma)} \delta_{k,n}, \quad \rho_n^{(\gamma)} = (p^n, p^n)_\gamma, \tag{26}$$

где $\gamma = 0, 1, 2$, а $\delta_{k,n}$ – символ Кронекера. При этих условиях для невязки имеем соотношения

$$\left(\bar{r}^{n+1}, \bar{r}^{n+1}\right)_{\gamma-2} = \left(\bar{r}^0, \bar{r}^0\right)_{\gamma-2} - \sum_{k=0}^{n} \left[2\alpha_k \left(p^k, p^k\right)_{\gamma-1} - \alpha_k^2 \left(p^k, p^k\right)_\gamma\right].$$

Отсюда получаем, что если коэффициенты α_n равны

$$\alpha_n = \sigma_n / \rho_n, \quad \sigma_k = \left(\bar{r}_0, p^k\right)_{\gamma-1}, \tag{27}$$

то для функционала невязки справедливы равенства

$$\Phi^\gamma\left(\bar{r}^{n+1}\right) = \Phi^\gamma\left(\bar{r}^0\right) - \sum_{k=0}^{n} \sigma_k^2 / \rho_k, \tag{28}$$

где знаки "γ" у α_k, σ_k и ρ_k для краткости опущены. При этом функционалы достигают минимума в подпространствах Крылова

$$\mathcal{K}_{n+1}\left(\bar{r}^0, \bar{A}\right) = \mathrm{Span}\left\{\bar{r}^0, \bar{A}\bar{r}^0, \ldots, \bar{A}^n\bar{r}^0\right\}. \tag{29}$$

Для обеспечения условий ортогональности (26) направляющие векторы могут определяться из рекурсий

$$p^0 = \bar{r}^0, \quad p^{n+1} = \bar{r}^{n+1} + \beta_n p_n, \quad \beta_n = -\left(\bar{r}^{n+1}, p^n\right)_\gamma / \rho_n, \tag{30}$$

где правило нахождения начального вектора p^0 является общепринятым, но не обязательным (строго говоря, он может быть произвольным). При этом выполняются дополнительные свойства ортогональности векторов

$$\left(r^k, r^n\right)_{\gamma-1} = ||r^n||_{\gamma-1}\delta_{k,n}, \quad \left(r^n, p^k\right)_{\gamma-1} = 0 \quad \text{при} \quad k < n,$$

а также равенства $(\bar{r}^0, p^n)_{\gamma-1} = (\bar{r}^n, \bar{r}^n)_{\gamma-1}$, из которых следуют новые формулы для коэффициентов σ_n, β_n:

$$\sigma_n = \left(\bar{r}^n, \bar{r}^n\right)_{\gamma-1}, \quad \beta_n = \sigma_{n+1}/\sigma_n. \tag{31}$$

Описанные алгоритмы сопряженных направлений при $\gamma = 1, 2$ имеют названия методов сопряженных градиентов и сопряженных невязок соответственно (обозначения CG и CR, от Conjugate Gradient и Conjugate Residual, см. [4, 14] и цитируемые там работы). Из приведенных соотношений для $\gamma = 1, 2$ получаем следующие формулы в терминах матриц A и B:

– для методов сопряженных градиентов:

$$\begin{gathered}
r^0 = f - Au^0, \quad p^0 = B^{-1}r^0, \quad \alpha_n = \sigma_n/\rho_n, \\
u^{n+1} = u^n + \alpha_n p_n, \quad r^{n+1} = r^n - \alpha_n A p^n, \quad p^{n+1} = B^{-1}r^{n+1} + \beta_n p^n, \\
\sigma_n = \left(B^{-1}r^n, r^n\right), \quad \rho_n = \left(Ar^n, r^n\right), \quad \beta_n = \sigma_{n+1}/\sigma_n;
\end{gathered} \tag{32}$$

– для методов сопряженных невязок:

$$\begin{gathered}
r^0 = f - Au^0, \quad \widehat{r}^0 = p^0 = B^{-1}r^0, \\
u^{n+1} = u^n + \alpha_n p_n, \quad \widehat{r}^{n+1} = \widehat{r}^n - \alpha_n B^{-1}Ap^n, \quad p^{n+1} = \widehat{r}^{n+1} + \beta_n p_n, \\
\sigma_n = \left(A\widehat{r}^n, \widehat{r}^n\right), \quad \rho_n = \left(B^{-1}Ap^n, Ap^n\right).
\end{gathered} \tag{33}$$

Отметим, что в обоих методах на каждой итерации требуется по одному умножению на матрицы A и B^{-1}, а в формулах (12) вектор $\widehat{r}^n$ – это не "настоящая", а предобусловленная невязка, т. е. при точных вычислениях имеем $\widehat{r}^n = B^{-1}r^n = B^{-1}(f - Au^n)$, а на каждой итерации при этом минимизируется величина $(B^{-1}r^n, r^n)$. Алгоритмы (32), (33) имеют распространенные обозначения PCG и PCR соответственно.

Для $\gamma = 0$ реализация алгоритма должна выполняться другим образом, поскольку нахождение σ_n в данном случае требует обращение матрицы $\bar{A}$. Получаемые при этом алгоритмы носят названия методов минимальных ошибок, или минимальных итераций. Первые их исследования приведены в [17–20], а более позднее, в связи с алгебраической проблемой моментов, – в [6]. Используя (4), запишем векторы ошибок и невязок в виде

$$\begin{aligned} v^n &= \bar{u} - \bar{u}^n = \alpha_n p^n + \cdots + \alpha_M p^M, \quad M \leqslant N, \\ \bar{r}^n &= A v^n = \alpha_n A p^m + \cdots + \alpha_M \bar{A} p^M. \end{aligned} \tag{34}$$

Умножая первые из этих соотношений скалярно на p^n, получаем

$$\alpha_n = (v^n, p^n)/(p^n, p^n) = -\alpha_{n-1}(\bar{A} v^n p^{n-1}) = -\alpha_{n-1}(\bar{r}^n, p^{n-1})/||p^n||.$$

Здесь использована следующая из (25), (30) трехчленная рекурсия, справедливая при $n \geqslant 1$ (для $n = 1$ $\beta_{-1} = 0$) :

$$p^n = (1 + \beta_{n-1}) p^{n-1} - \alpha_{n-1} \bar{A} p^{n-1} - \beta_{n-2} P^{n-2},$$

а также ортогональность направляющих векторов p^n и симметричность матрицы $\bar{A}$. При $n = 0$ имеем

$$\alpha_0 = (v^0, p^0)/(p^0, p^0).$$

Отсюда, в силу произвольности начального направляющего вектора, мы можем положить $p^0 = \bar{A}\bar{r}^0$, откуда получаем

$$\alpha_0 = (\bar{r}^0, \bar{r}^0)/(\bar{A}\bar{r}^0, \bar{A}\bar{r}^0) = (B^{-1} r^0, r^0)/(B^{-1} A B^{-1} r^0, A B^{-1} r^0). \tag{35}$$

Коэффициенты β_n при этом вычисляются по формуле (30).

Один из актуальных вопросов при реализации методов (11) – (12) возникает, когда обращение предобуславливающей матрицы B выполняется приближенно и фактически сводится к итерационному решению, соответствующей вспомогательной СЛАУ. При этом допускается ненулевая конечная невязка и ошибка. Сам итерационный процесс в этом случае яввляется двухуровневым. Хотя вопрос о критериях окончания итераций достаточно тонкий, мы для простоты считаем, что в предобусловленных алгоритмах CG и CR внешний процесс завершается при выполнении условия

$$||r^n|| \leqslant \varepsilon_e ||f||, \quad \varepsilon_e \ll 1. \tag{36}$$

На внутренних итерациях поступаем аналогично, выбирая некоторый другой параметр точности $\varepsilon_i \leqslant 1$. Например, приближенный начальный направляющий вектор $\bar{p}^0$ определяется из решения системы $\bar{B}p^0 = r^0$ следующим образом:

$$\delta^0 = r^0 - B\bar{p}^0, \quad \bar{p}^0 = B^{-1}(r^0 - \delta^0), \quad ||\delta^0|| \leqslant \varepsilon_i ||r^0||. \tag{37}$$

§5. Заключение

Рассмотренные подходы к построению посменно треугольных предобуславливающих матриц открывают новые возможности для распараллеливания простых и/или блочных итерационных методов симметричной последовательной верхней релаксации, а также неполной факторизации в подпространствах Крылова (в том числе вложенной). Пока что открытым остается теоретический вопрос о скорости сходимости получаемых новых итерационных процессов, так что здесь требуется, в первую очередь, провести систематические экспериментальные исследования с апробацией различных модификаций алгоритмов. Конечной целью в данном случае является ускорение и повышение производительности вычислений на многопроцессорных системах с распределенной и иерархической общей памятью.

Список литературы

1. В. П. Ильин, *Математическое моделирование. Часть 1. Непрерывные и дискретные модели.* Изд. СО РАН, Новосибирск, 2017.
2. O. Axelsson, *Iterative Solution Methods.* Cambridge Univ. Press, Cambridge, 1994.
3. В. П. Ильин, *Методы неполной факторизации для решения алгебраических систем.* Наука, М., 1995.
4. Y. Saad, *Iterative Methods for Sparse Linear Systems. 2nd edn.* SIAM, 2003.
5. В. П. Ильин, *Методы и технологии конечных элементов.* ИВМиМГ СО РАН, Новосибирск, 2007.
6. J. Liesen, Z. Strakos, *Krylov Subspace Methods, Principles and Analysis.* Oxford University Press, 2013.
7. M. A. Olshanskii, E. E. Tyrtyshnikov, *Iterative Methods for Linear Systems Theory and Applications.* SIAM, Philadelphia, 2014.
8. V. P. Il'in, K. Yu. Laevsky, *Generalized compensation principle in incomplete factorization methods.* — Russ. J. Numer. Anal. Math. Model. **12**, No. 5 (1997), 399–412.
9. R. Wang, Q. Niu, L. Lu, *A weighted block tangential filtering decomposition preconditioner.* — Math. Prob. Engineering (2009). Article ID 282307.
10. В. П. Ильин, Ю. И. Кузнецов, *Трехдиагональные матрицы и их приложения.* Наука, М., 1985.

11. J. R. Appleyard, I. M. Cheshire, R. K. Pollard, *Special techniques for fully-implicit simulators.* — Proc. European Symposium on Enhanced Oil Recovery Bournemouth, 1981, 395–408.
12. J. R. Appleyard, I. M. Cheshire, *Nested factorization.* — Proc. Seventh SPE Symposium on Reservoir Simulation, 1983, 315–324.
13. N. N. Kuznetsova, O. V. Diyankov, S. V. Kotegov, I. V. Krasnogorov, V. Y. Pravilnikov, S. Y. Maliassov, *The family of nested factorizations.* — RJNAMM **22** (2007), 393–412.
14. P. Kumar, L. Grigori, F. Nataf, Q. Nui, *On relaxed nested factorization and combination preconditioning.* — Int. J. Comput. Math. **93**, No. 1 (2016), 178–199.
15. P. Kumar, L. Grigori, Q. Niu, F. Nataf, *Fourier analysis of modified nested factorization preconditioner for three-dimensional isotropic problems.* Res. Rep. 2010. INRIA - 00448291.
16. *Using ECLIPSE Parallel.* Parallel computing, 2003. `www.sis.slb.com/me-dia/about/whitepaper_parallelcomputing.pdf.`
17. E. Graig, *The n-steps iterative procedures.* — J. Math. Phys. **34** (1955), 64–73.
18. В. М. Фридман, *Метод минимальных итераций с минимальными ошибками для системы линейных алгебраических уравнений с симметричной матрицей.* — Ж. вычисл. матем. и матем. физ. **2**, No. 2 (1962), 341–342.
19. Ю. В. Воробьев, *Метод моментом и прикладной математике.* Физматлит, М., 1958.
20. Д. К. Фаддеев, В. Н. Фаддеева, *Вычислительные методы линейной алгебры.* Физматгиз, М., 1960.
21. Y. Saad, M. Yeung, J. Erhel, F. Guyomarc'h, *A deflated version of conjugate gradient algorithm.* — SIAM J. Sci. Comput. **24** (2000), 1909–1926.
22. Л. Ю. Колотилина, *Переобуславливание систем линейных алгебраических уравнений с помощью двойного исчерпывания.* I. *Теория.* — Зап. научн. семин. ПОМИ **229** (1995), 95–152.
23. N. Venkovic, P. Mycek, L. Giraud, O. Le Maitre, *Comparative study of harmonic and Rayleigh–Ritz procedures with applications for deflated conjugate gradients.* Pressed Report Cerfacs, 2020. hal-02434043

ll'in V. P. Parallel changing triangular iterative methods in Krylov subspaces.

The paper considers parallel preconditioned iterative methods in Krylov subspaces for solving large systems of linear algebraic equations with sparse symmetric positive-definite matrices arising in grid approximations of multidimensional problems. For preconditioning, generalized block algorithms of symmetric successive overrelaxation or incomplete factorization with matching row sums are used. Preconditioners are based on changing triangular matrix factors with multiple changes in the triangulation structure. In three-dimensional grid algebraic systems, methods are based on nested factorizations, as well as on two-level iterative processes. Successive approximations in Krylov subspaces are computed by applying a family of conjugate direction algorithms with various orthogonal and variational properties, including preconditioned conjugate gradient methods, conjugate residuals, and minimal errors.

Институт вычислительной математики
и математической геофизики СО РАН
Новосибирский государственный
университет, Новосибирск, Россия
E-mail: ilin@sscc.ru

Поступило 23 октября 2020 г.