

МЕТОД ПАРАЛЛЕЛЬНОЙ МАТРИЧНОЙ ПРОГОНКИ С УМЕНЬШЕННОЙ РЕДУЦИРОВАННОЙ СИСТЕМОЙ

С.В. Баханович¹, Н.А. Лиходед²

¹*Институт математики НАН Беларуси,
ул. Сурганова, 11, 220072, г. Минск, Беларусь, bsv@im.bas-net.by*

²*Белорусский государственный университет,
пр. Независимости, 4, 220030 г. Минск, Беларусь, likhoded@bsu.by*

Предложена новая модификация алгоритма параллельной матричной прогонки. Основным отличием предлагаемой версии от известных параллельных алгоритмов матричной прогонки является наличие в возникающей редуцированной системе вдвое меньшего числа матрично-векторных уравнений. Если для исходной системы уравнений справедливы условия устойчивости метода матричной прогонки, то вычисления разработанной модификации также являются устойчивыми.

Ключевые слова: параллельные вычисления; матричная прогонка; устойчивость параллельной матричной прогонки.

THE METHOD OF PARALLEL MATRIX SWEEP WITH A DECREASED REDUCED SYSTEM

S.V. Bakhanovich^a, N.A. Likhoded^b

^a*Institute of Mathematics, National Academy of Science,
6 Surhanava Street, Minsk 220072, Belarus, bsv@im.bas-net.by*

^b*Belarusian State University,
4 Niezalieznasci Avenue, Minsk 220030, Belarus, likhoded@bsu.by*

A new modification of the parallel matrix sweep is proposed. The main difference of the proposed version from the well-known parallel matrix sweep algorithms is the presence in the resulting reduced system of half the number of matrix-vector equations. If the stability conditions of the matrix run method are valid for the initial system of equations, then the calculations of the developed modification are also stable.

Keywords: parallel computing; matrix sweep; stability of parallel matrix sweep.

Введение

Метод матричной прогонки [1] широко применяется для решения систем линейных алгебраических уравнений с блочно-трехдиагональными матрицами. Такие системы часто возникают в практических приложениях при численном решении уравнений в частных производных. Размерность систем и их блочных коэффициентов в ряде

случаев не позволяет использовать последовательные алгоритмы, поэтому нужны параллельные алгоритмы матричной прогонки, эффективные для реализации на многоядерных вычислительных системах.

Известны параллельные алгоритмы матричной прогонки, основанные на разбиении матрицы системы [2, 3]. Эти алгоритмы обобщают на случай блочно-трехдиагональных матриц идеи скалярных алгоритмов параллельной прогонки [4, 5]. Методы, основанные на разбиении матрицы, являются трехфазными. На первой фазе исходная система разбивается на части и после независимых преобразований каждой части составляется редуцированная система (блочно-трехдиагональная в блочном случае). На второй – решая редуцированную систему, находят несколько неизвестных каждой части уравнений. На третьей фазе независимо вычисляются остальные неизвестные каждой части.

В этой работе предложена новая модификация параллельной матричной прогонки, которая также является обобщением на блочный случай скалярного алгоритма параллельной прогонки [6]. Основным отличием предлагаемой версии от известных параллельных алгоритмов матричной прогонки является наличие в редуцированной системе вдвое меньшего числа матрично-векторных уравнений. Обосновано, что если для исходной системы уравнений справедливы известные условия устойчивости метода матричной прогонки, то вычисления разработанной модификации параллельной матричной прогонки являются устойчивыми.

1. Матричная прогонка

Пусть $A_1, \dots, A_N, B_0, B_1, \dots, B_{N-1}, C_0, C_1, \dots, C_N$ – $M \times M$ -матрицы, $Y_0, Y_1, \dots, Y_N, F_0, F_1, \dots, F_N$ – M -мерные векторы. Рассмотрим блочно-трехдиагональную систему линейных алгебраических уравнений вида

$$C_0 Y_0 - B_0 Y_1 = F_0, \quad -A_i Y_{i-1} + C_i Y_i - B_i Y_{i+1} = F_i, \quad i = 1, \dots, N-1, \quad -A_N Y_{N-1} + C_N Y_N = F_N.$$

Приведем формулы матричной прогонки для решения исходной системы [1]:

- прямой ход метода прогонки – вычисление матриц α_i и векторов β_i по формулам

$$\alpha_1 = C_0^{-1} B_0, \quad \beta_1 = C_0^{-1} f_0,$$

$$\alpha_{i+1} = (C_i - A_i \alpha_i)^{-1} B_i, \quad \beta_{i+1} = (C_i - A_i \alpha_i)^{-1} (F_i + A_i \beta_i), \quad i = 1, 2, \dots, N-1,$$

$$\beta_{i+1} = (C_i - A_i \alpha_i)^{-1} (F_i + A_i \beta_i);$$

- обратный ход прогонки – вычисление решения по формулам $Y_i = \beta_{i+1}, Y_i = \alpha_{i+1} Y_{i+1} + \beta_{i+1}, i = N-1, \dots, 1, 0.$

2. Параллельная матричная прогонка, основанная на разбиении блочной трехдиагональной матрицы

Пусть дана блочно-трехдиагональная система. Разобьем эту систему, состоящую из $N+1$ матрично-векторных уравнений, на K частей, т.е. на K блоков уравнений, где K – параметр алгоритма. Уравнения в системе будем нумеровать, начиная с нуля; для простоты изложения будем считать, что $N+1$ нацело делится на K : $\frac{N+1}{K}=l$, где l – целое число. В k -й блок уравнений, $k=0,1,\dots,K-1$, входят l матрично-векторных уравнений, начиная с уравнения $k \cdot l$ и по уравнение $(k+1)l-1$ включительно. Обозначим $s_k=k \cdot l$ номер первого уравнения в k -м блоке, $f_k=(k+1)l-1$ – номер последнего уравнения в k -м блоке.

Независимо решить блоки уравнений не представляется возможным, так как первое и последнее уравнения блока содержит переменные, которые входят в соседние блоки. Будем называть такие переменные граничными.

Алгоритмы параллельной матричной прогонки состоят из трех фаз. На первой фазе уравнения каждого из K блоков независимо преобразуются так, чтобы можно было составить так называемую редуцированную блочно-трехдиагональную систему, содержащую только граничные неизвестные, а остальные неизвестные можно было найти, зная граничные неизвестные. На второй фазе решается редуцированная система (последовательно или параллельно). После чего на третьей фазе, используя найденные на второй фазе неизвестные, независимо вычисляются оставшиеся неизвестные.

Основным отличием предлагаемой версии от известных параллельных алгоритмов матричной прогонки является наличие в редуцированной системе K , а не $2K$ матрично-векторных уравнений.

Пусть произведено разбиение системы на K блоков уравнений. В каждом из блоков, кроме нулевого, в первом уравнении есть ненулевые, вообще говоря, коэффициенты при неизвестных с индексами s_{k-1} и s_k ; эти же неизвестные входят также в уравнения предыдущего блока. Кроме того, в каждом блоке, кроме $(K-1)$ -го, в последнем уравнении есть ненулевые коэффициенты при неизвестных с индексами s_k+l-1 и s_k+l , которые входят также в уравнения следующего блока. Это не позволяет решать блоки уравнений независимо.

Неизвестные с индексами s_{k-1} , s_k и s_k+l-1 , s_k+l , где k – номера внутренних блоков являются граничными. Неизвестные последнего уравнения системы также отнесем к граничным. Рассмотрим способ получения редуцированной системы, в которую для каждого k войдет уравнение отно-

сительно граничных неизвестных k -го блока уравнений Y_{s_k-1} (кроме $k=0$), Y_{s_k+l-1} , и граничного неизвестного $(k+1)$ -го блока уравнений Y_{s_k+2l-1} (кроме $k=K-1$). Граничные неизвестные k -го блока Y_{s_k} и Y_{s_k+l} в редуцированную систему не входят.

Прямой ход первой фазы вычислений состоит в исключении строчными преобразованиями поддиагональных элементов в матрице каждого блока уравнений. В каждом k -м блоке исключаются все поддиагональные элементы, кроме $-A_{s_k}$ ($k \neq 0$) в первом уравнении блока (в уравнении с номером s_k исходной системы). При этом, кроме случая $k \neq 0$, возникнет столбец ненулевых коэффициентов при неизвестном Y_{s_k-1} , т.е. при неизвестном $Y_{f_{k-1}}$.

На шаге r , $r=s_k+1, \dots, f_k$, преобразуется уравнение исходной системы относительно неизвестных с индексами $r-1, r, r+1$. Преобразованные коэффициенты-блоки будем хранить в массивах L_r^A (при неизвестном $Y_{s_k-1} = Y_{f_{k-1}}$), L_r^C (при неизвестных Y_r), L_r^F (правая часть уравнения с номером r); коэффициенты-блоки при неизвестных Y_{r+1} есть $-B_{r+1}$. В случае $k=0$ столбец, содержащий L_{r-1}^A , отсутствует.

После шага $r=f_k$ последнее уравнение в блоке становится уравнением относительно неизвестных с индексами s_k-1, s_k+l-1, s_k+l , т.е. относительно неизвестных с индексами f_{k-1}, f_k, f_{k+1} , $k=0, 1, \dots, K-1$. Новые массивы L_r^A, L_r^C, L_r^F будут использоваться при выполнении обратного хода первой фазы вычислений и (частично) на второй и третьей фазах.

Обратный ход первой фазы вычислений состоит в исключении строчными преобразованиями наддиагональных элементов в матрице каждого блока уравнений. Исключаются все наддиагональные элементы, кроме $-B_{f_k-1}$ в предпоследнем уравнении блока (в уравнении с номером f_k-1 исходной системы). При этом в каждом k -м блоке возникнет столбец ненулевых коэффициентов при неизвестном Y_{s_k+l-1} , т.е. при неизвестном Y_{f_k} . Возникнет также ненулевой коэффициент при неизвестном Y_{s_k+2l-1} , т.е. при неизвестном $Y_{f_{k+1}}$, в последнем уравнении k -го блока (кроме случая $k=K-1$). Это происходит потому, что с помощью преобразованной при обратном ходе первой строки $(k+1)$ -го блока обнуляется наддиагональный элемент в последнем уравнении k -го блока.

Из последних уравнений блока, т.е. из уравнений с номером f_k , $k=0,1,\dots,K-1$, исходной системы, составляется редуцированная блочно-трехдиагональная система относительно неизвестных Z_{k-1} ($k \neq 0$), Z_k , Z_{k+1} ($k \neq K-1$), где $Z_{k-1} = Y_{f_{k-1}}$, $Z_k = Y_{f_k}$, $Z_{k+1} = Y_{f_{k+1}}$.

Рассмотрим вторую фазу вычислений. Обозначим $-\bar{A}_k = L_{f_k}^A$, $k=1,\dots,K-1$; $\bar{C}_k = \bar{L}_{f_k}^C$, $k=0,1,\dots,K-2$; $-\bar{B}_k = L_{f_k}^B$, $k=0,1,\dots,K-2$; $\bar{F}_k = \bar{L}_{f_k}^F$, $k=0,1,\dots,K-2$, $\bar{F}_{K-1} = \bar{L}_N^F$. Тогда, с учетом введенного обозначения $Z_k = Y_{f_k}$ эта система примет вид

$$\begin{aligned}\bar{C}_0 Z_0 - \bar{B}_0 Z_1 &= \bar{F}_0 \quad (k=0), \\ -\bar{A}_k Z_{k-1} + \bar{C}_k Z_k - \bar{B}_k Z_{k+1} &= \bar{F}_k, \quad k=1,\dots,K-2, \\ -\bar{A}_{K-1} Z_{K-2} + \bar{C}_{K-1} Z_{K-1} &= \bar{F}_{K-1} \quad (k=K-1).\end{aligned}$$

На этапе второй фазы решается (например, методом матричной прогонки) полученная относительно $Z_0, Z_1, \dots, Z_{K-1}$ система. Одновременно с этим можно в качестве предварительного этапа третьей фазы вычислений для каждого из K блоков независимо получать $(L_r^C)^{-1}$, $r=s_k, \dots, f_k-1$.

На второй фазе алгоритма найдены $Y_{f_0}, Y_{f_1}, \dots, Y_{f_{K-1}}$. На третьей фазе алгоритма для каждого блока уравнений k , $k=0,1,\dots,K-1$, используя найденные на второй фазе неизвестные, вычисляются оставшиеся неизвестные Y_r , $r=s_k, \dots, f_k-1$.

Рассмотрим в блоке k преобразованные на первой фазе уравнения, не вошедшие в редуцированную систему. Перенесем в каждом уравнении с номером r известные произведения $\bar{L}_r^A Y_{f_{k-1}}$ (кроме случая $k=0$) и $L_r^B Y_{f_k}$ в правую часть. Тогда независимо для каждого из блоков можно найти Y_r :

$$Y_r = (L_r^C)^{-1} \left(\bar{L}_{f_k}^F - \bar{L}_r^A Y_{f_{k-1}} - L_r^B Y_{f_k} \right), \quad r=s_k, \dots, f_k-1.$$

3. Устойчивость параллельной матричной прогонки с уменьшенной редуцированной системой

Говорят, что алгоритм метода матричной прогонки устойчив, если выполнено условие $\|\alpha_i\| \leq 1$ для $i=1,2,\dots,N$ (предполагается, что в M -мерном пространстве введена какая-либо норма). Имеет место следующее утверждение [1]: если C_i , $i=0,1,\dots,N$, – невырожденные матрицы, а $A_1,\dots,A_N$, $B_0,B_1,\dots,B_{N-1}$ – ненулевые матрицы и выполнены условия $\|C_i^{-1}A_i\| + \|C_i^{-1}B_i\| \leq 1$, $i=1,\dots,N-1$, $\|C_0^{-1}B_0\| \leq 1$, $\|C_N^{-1}A_N\| \leq 1$, причем хотя бы в одном из этих условий выполняется строгое неравенство, то алгоритм метода матричной прогонки устойчив.

Теорема. Если для исходной блочно-трехдиагональной системы выполнены условия устойчивости метода последовательной матричной прогонки причем хотя бы для одного уравнения, не являющегося последним в блоке, неравенство $\|C_i^{-1}A_i\| + \|C_i^{-1}B_i\| \leq 1$ является строгим, то условия устойчивости выполняются для редуцированной системы и для всех уравнений на третьей фазе вычислений.

Таким образом, разработан и исследован с точки зрения устойчивости новый вариант метода параллельной матричной прогонки для решения систем линейных алгебраических уравнений с блочно-трехдиагональными матрицами.

Библиографические ссылки

1. Самарский А.А. Методы решения сеточных уравнений. М.: Наука. 1978. 592 с.
2. Акимова Е.Н. Параллельные алгоритмы решения СЛАУ с блочно-трехдиагональными матрицами на многопроцессорных вычислителях // Вестн. УГАТУ. 2011. № 15(5). С. 87–93.
3. Згировский А.А., Лиходед Н.А. Модифицированный метод параллельной матричной прогонки // Вес. Нац. акад. наук Беларусі. Сер. фіз.-мат. навук. 2019. № 55(4). С. 425–434. DOI: 10.29235/1561-2430-2019-55-4-425-434.
4. Яненко Н.Н., Коновалов А.Н., Бугров А.Н., Шустов Г.В. Об организации параллельных вычислений и “распараллеливании” прогонки // Численные методы механики сплошной среды. 1978. № 9(7). С. 139–146.
5. Austin T.M., Berndt M., Moulton J.D. A memory efficient parallel tridiagonal solver. Preprint LA-VR-03-4149. 2004. 13 p.
6. Баркалов К.А. Методы параллельных вычислений. Н. Новгород: Изд-во Нижегородского госуниверситета им. Н.И. Лобачевского. 2011. 124 с.