О коэффициенте при логарифме в критическом пути графа циклической редукции

А.В. Фролов[1]

ИВМ РАН[1]

Автор разбирает альтернативный наиболее распространённому вариант метода циклической редукции. На его взгляд, у нового варианта преимущества по сравнению с традиционным не только в более коротком критическом пути, но и в более широком диапазоне размерностей решаемых трехдиагональных систем линейных алгебраических уравнений[1].

Ключевые слова: циклическая редукция, критический путь, граф алгоритма.

1. Введение

Задача эффективного решения трёхдиагональных СЛАУ давно вызывает интерес исследователей [5]. Несмотря на отклики авторитетных исследований об её устарении как излишне лёгкой и эффективно реализуемой на однопроцессорных устройствах, отклики прикладников, столдкнувшихся с такими задачами, на [10] и непосредственное исследование производительности в [9] показывают, что эта задача далеко не исчерпала себя и по-прежнему остаётся интересной. В данной статье автор, уже обращавший на неё внимание читателей в [7–10], предлагает по-новому посмотреть на такой способ решения системы линейных алгебраических уравнений с трёхдиагональными и блочно-трёхдиагональными матрицами, как метод циклической редукции. Традиционно [5] этот метод реализуют схемой, напоминающей сдваивание, для размерностей неизвестных векторов на 1 меньше, чем степень двойки, поскольку именно для таких размерностей эта схема максимально проста и при ней достижима теоретически оценка пути критического графа алгоритма. Автор задался вопросом: а нельзя ли сходным методом с чуть более сложной схемой улучшить оценку критического пути, попутно обобщив её на гораздо больший спектр размерностей решаемых трехдиагональных СЛАУ. Как выяснилось, ответ на этот вопрос положителен, что будет показано ниже.

2. Циклическая редукция и её модификация

Сначала вспомним классическую схему циклической редукции и оценим её характеристики, а потом рассмотрим новую схему. Решаемую СЛАУ будем считать имеющей вид

$$
\begin{gathered}
b_1x_1 + c_1x_2 = f_1 \\
a_2x_1 + b_2x_2 + c_2x_3 = f_2 \\
\dots \\
a_ix_{i-1} + b_ix_i + c_ix_{i+1} = f_i \\
\dots \\
a_{n-1}x_{n-2} + b_{n-1}x_{n-1} + c_{n-1}x_n = f_{n-1} \\
a_nx_{n-1} + b_nx_n = f_n
\end{gathered}
$$

Здесь $n = 2^k - 1$.

2.1 Классическая схема со степенями двойки

Для редукции исходной СЛАУ к СЛАУ с размерностью $2^{k-1} - 1$ применяется подстановка из уравнений с нечётными номерами в уравнения с чётными. Пусть i чётно. Тогда из *i-1*-го уравнения

[1] далее для систем линейных алгебраических уравнений в статье будем использовать сокращение СЛАУ.

$$a_{i-1}x_{i-2} + b_{i-1}x_{i-1} + c_{i-1}x_i = f_{i-1} \quad (1)$$

получаем

$$x_{i-1} = \frac{f_{i-1}}{b_{i-1}} - \frac{a_{i-1}}{b_{i-1}} x_{i-2} - \frac{c_{i-1}}{b_{i-1}} x_i \quad (2)$$

Аналогично, из *i+1*-го уравнения

$$a_{i+1}x_i + b_{i+1}x_{i+1} + c_{i+1}x_{i+2} = f_{i+1} \quad (3)$$

$$x_{i+1} = \frac{f_{i+1}}{b_{i+1}} - \frac{a_{i+1}}{b_{i+1}} x_i - \frac{c_{i+1}}{b_{i+1}} x_{i+2} \quad (4)$$

После подстановки найденных выражений в *i*-е уравнение

$$a_i x_{i-1} + b_i x_i + c_i x_{i+1} = f_i \quad (5)$$

получаем уравнение, связывающее соседей с чётными номерами:

$$A_i x_{i-2} + B_i x_i + C_i x_{i+2} = F_i \quad (6)$$

где

$$A_i = -\frac{a_{i-1}}{b_{i-1}} a_i,\ B_i = b_i - \frac{c_{i-1}}{b_{i-1}} a_i - \frac{a_{i+1}}{b_{i+1}} c_i,\ C_i = -\frac{c_{i+1}}{b_{i+1}} c_i, \qquad F_i = f_i - \frac{f_{i-1}}{b_{i-1}} a_i - \frac{f_{i+1}}{b_{i+1}} c_i \quad (7)$$

(при *i=2* в формулах (7) будут отсутствовать члены с a_i, при *i=n-1* – с c_{i+1}).

Последовательным циклическим применением такой редукции оставляем в СЛАУ только компоненты с номерами, делящимися на 2, потом на 4, на 8 и т.д., пока не получаем уравнение с компонентом, имеющим номер $\frac{n+1}{2}$. Решаем его, деля правую часть на коэффициент, после чего подставляем уже полученные компоненты в уравнения, аналогичные (2) и (4), на каждом шаге удваивая количество полученных компонентов. Решив все уравнения (2) и (4), заканчиваем решение всей СЛАУ.

Оценим критический путь графа алгоритма циклической редукции. Приём редукции применяется *k-1* раз, при этом все вычисления в формулах (7) можно сделать параллельно, предварительно вычислив все отношения, которые встречаются там, одновременно зафиксировав их в формулах (2) и (4). То есть на каждом из *k-1* шагов редукции у нас вначале будет ярус с делениями, а потом ярус с умножениями и два яруса с вычитаниями/сложениями. Потом идёт одно деление и обратные *k-1* шагов (формулы (2) и (4)), на каждом из них по ярусу умножения и по два яруса вычитания/сложения. Подводим итог по полной классической редукции: *k* ярусов делений, 2*k-2* умножений и 4*k-4* вычитаний/сложений. Если же у нас редукция проводится повторно (с той же матрицей СЛАУ, но с новой правой частью), то из отношений в (2) и (4) вычисляются только самые первые, а из формул (7) – только последняя. Получаем (учитывая, что деление можно заменить умножением на предвычисленную обратную величину) *3k-2* ярусов умножений и *4k-4* вычитаний/сложений. Остаётся перевести формулы критического пути к виду, зависимому от *n*. Получаем, что у нас $3\lceil log_2(n+1)\rceil - 2$ умножений и $4\lceil log_2(n+1)\rceil - 4$ вычитаний/сложений.

2.2 Новая схема

Присмотримся к формулам (2), (4) и (7). Во всех них по 2 вычитания, и выполняются они с необходимостью на 2 ярусах, непараллельно. Можно ли исправить положение? Да! Выполним редукцию асимметрично. Будем оставлять те неизвестные, номера которых дают в остатке от деления на 5 ноль или 3. Пусть сначала *i=5m+3*.

Из

$$a_{i-2}x_{i-3} + b_{i-2}x_{i-2} + c_{i-2}x_{i-1} = f_{i-2} \quad (8)$$

и

$$a_{i-1}x_{i-2} + b_{i-1}x_{i-1} + c_{i-1}x_i = f_{i-1} \quad (1)$$

по правилу Крамера выразим x_{i-1}:

$$x_{i-1} = \frac{\begin{vmatrix} b_{i-2} & f_{i-2} - a_{i-2}x_{i-3} \\ a_{i-1} & f_{i-1} - c_{i-1}x_i \end{vmatrix}}{\begin{vmatrix} b_{i-2} & c_{i-2} \\ a_{i-1} & b_{i-1} \end{vmatrix}} = \frac{b_{i-2}(f_{i-1} - c_{i-1}x_i) - a_{i-1}(f_{i-2} - a_{i-2}x_{i-3})}{b_{i-1}b_{i-2} - a_{i-1}c_{i-2}}$$

$$x_{i-1} = \frac{b_{i-2}}{b_{i-1}b_{i-2} - a_{i-1}c_{i-2}} f_{i-1} - \frac{a_{i-1}}{b_{i-1}b_{i-2} - a_{i-1}c_{i-2}} f_{i-2} - \frac{b_{i-2}c_{i-1}}{b_{i-1}b_{i-2} - a_{i-1}c_{i-2}} x_i + \frac{a_{i-1}a_{i-2}}{b_{i-1}b_{i-2} - a_{i-1}c_{i-2}} x_{i-3} \quad (9)$$

Из (9) и (4) после подстановки в (5) получаем

$$A_i x_{i-3} + B_i x_i + C_i x_{i+2} = F_i \quad (10)$$

где

$$A_i = \frac{a_i a_{i-1} a_{i-2}}{b_{i-1}b_{i-2} - a_{i-1}c_{i-2}}, \quad C_i = -\frac{c_{i+1}}{b_{i+1}} c_i \quad (11)$$

$$B_i = b_i - \frac{b_{i-2}c_{i-1}}{b_{i-1}b_{i-2} - a_{i-1}c_{i-2}} a_i - \frac{a_{i+1}}{b_{i+1}} c_i \quad (12)$$

$$F_i = f_i - \frac{b_{i-2}a_i}{b_{i-1}b_{i-2} - a_{i-1}c_{i-2}} f_{i-1} + \frac{a_{i-1}a_i}{b_{i-1}b_{i-2} - a_{i-1}c_{i-2}} f_{i-2} - \frac{f_{i+1}}{b_{i+1}} c_i \quad (13)$$

(при *i=3* в формулах (10-13) будут отсутствовать члены с a_{i-2}, при *i=n-1* – с c_{i+1}).

Для $i=5m$ формулы аналогичны, но исключаются 1 неизвестное "снизу" и 2 "сверху": из

$$a_{i+1}x_i + b_{i+1}x_{i+1} + c_{i+1}x_{i+2} = f_{i+1} \quad (3)$$

и

$$a_{i+2}x_{i+1} + b_{i+2}x_{i+2} + c_{i+2}x_{i+3} = f_{i+2} \quad (14)$$

$$x_{i+1} = \frac{\begin{vmatrix} f_{i+1} - a_{i+1}x_i & c_{i+1} \\ f_{i+2} - c_{i+2}x_{i+3} & b_{i+2} \end{vmatrix}}{\begin{vmatrix} b_{i+1} & c_{i+1} \\ a_{i+2} & b_{i+2} \end{vmatrix}} = \frac{b_{i+2}(f_{i+1} - a_{i+1}x_i) - c_{i+1}(f_{i+2} - c_{i+2}x_{i+3})}{b_{i+1}b_{i+2} - a_{i+2}c_{i+1}}$$

$$x_{i+1} = \frac{b_{i+2}}{b_{i+1}b_{i+2} - a_{i+2}c_{i+1}} f_{i+1} - \frac{c_{i+1}}{b_{i+1}b_{i+2} - a_{i+2}c_{i+1}} f_{i+2} - \frac{b_{i+2}a_{i+1}}{b_{i+1}b_{i+2} - a_{i+2}c_{i+1}} x_i + \frac{c_{i+1}c_{i+2}}{b_{i+1}b_{i+2} - a_{i+2}c_{i+1}} x_{i+3} \quad (15)$$

Из (1) и (15) после подстановки в (5) получаем

.

$$A_i x_{i-2} + B_i x_i + C_i x_{i+3} = F_i \quad (16)$$

где

$$A_i = -\frac{a_{i-1}}{b_{i-1}} a_i, \quad C_i = \frac{c_i c_{i+1} c_{i+2}}{b_{i+1}b_{i+2} - a_{i+2}c_{i+1}} \quad (17)$$

$$B_i = b_i - \frac{c_{i-1}}{b_{i-1}} a_i - \frac{b_{i+2}a_{i+1}}{b_{i+1}b_{i+2} - a_{i+2}c_{i+1}} c_i \quad (18)$$

$$F_i = f_i - \frac{f_{i-1}}{b_{i-1}} a_i - \frac{b_{i+2}c_i}{b_{i+1}b_{i+2} - a_{i+2}c_{i+1}} f_{i+1} - \frac{c_{i+1}c_i}{b_{i+1}b_{i+2} - a_{i+2}c_{i+1}} f_{i+2} \quad (19)$$

(при *i=n-2* в формулах (10-13) будут отсутствовать члены с c_{i+2}).

При проведении редукции таким образом количество остающихся переменных и уравнений каждый раз будет уменьшаться не менее, чем в 5/2 раз[1]. Поэтому количество шагов (прямых и обратных) может быть оценено как $\lceil log_{5/2} n \rceil \approx \lceil \frac{3}{4} log_2 n \rceil$, что меньше, чем при стандартной циклической редукции. Остаётся только подсчитать критические пути локальных микрографов для формул (9), (11-13), (15), (17-19). Оказывается, что, поскольку все коэффициенты для формул (9) и (15), дающих обратных ход редукции, вычисляются сразу при прямом ходе, то для шагов обратной редукции оценка остаётся той же, что и в классической схеме, на каждом из них по ярусу умножения и по два яруса вычитания/сложения. Для прямого же хода, при внимательном подсчёте, вместо имеющихся в стандартной схеме яруса с делениями, яруса с

[1] На самом деле при сравнительно небольших размерностях СЛАУ редуцируется с ещё большим коэффициентом. Для иллюстрации этого можно привести обратную оптимальную последовательность размерностей: 1, 4, 12, 32, 82, 207, 519, 1299, ...

умножениями и двух ярусов с вычитаниями/сложениями, будет замена вычитаний на чуть более долгие умножения. Таким образом, можно считать, что прямой ход такой модифицированной циклической редукции будет[1] не медленнее прямого хода стандартной, а обратный - на 25% быстрее.

Ещё лучше новая схема в сравнении с классической выглядит, если рассматривать повторную циклическую редукцию для новой схемы. Все шаги тогда будут иметь столько же ярусов, как в стандартной циклической редукции, но их будет в 4/3 раза меньше.

Однако более важным результатом, по мнению автора, является большая универсальность предложенной схемы в отношении размерностей СЛАУ. Действительно, чередующееся разбиение по 3-2 уравнения может быть выполнено, с её обрыванием, практически для любой размерности. Возможно, что на основе этих схем будут написаны программы циклической редукции в тех областях, прикладников которых до сих пор удерживали от этого неподходящие для них размерности степени двойки.

3. Заключение

Автору в данной статье удалось показать, что коэффициент при двоичном логарифме в оценке критического пути графа алгоритма вполне улучшаем. Для повторного решения СЛАУ с той же матрицей и новой правой частью коэффициент ускорения достигает 4/3. Кроме этого, модифицированная схема, хоть и сложнее, но более гибка по отношению к размерности решаемых систем.

Вместе с тем, автор напоминает, что новая схема циклической редукции в полной мере наследует такие её черты, как неполная загрузка устройств, избыточность вычислений в сравнении с классической прогонкой и т.п. описанные в [9, 10].

Литература

1. Антонов А.С., Воеводин Вад.В., Воеводин Вл.В., Теплов А.М., Фролов А.В.. Первая версия Открытой энциклопедии свойств алгоритмов // Вестник УГАТУ. Серия управление, вычислительная техника и информатика. Том 19, N 2(68), 2015., С.150-159
2. Воеводин В.В. Вычислительные основы линейной алгебры. М.: Наука, 1977.
3. Воеводин В.В., Кузнецов Ю.А. Матрицы и вычисления. М.: Наука, 1984.
4. Воеводин В.В. Математические основы параллельных вычислений// М.: Изд. Моск. ун-та, 1991. 345 с.
5. Ильин В.П., Кузнецов Ю.И. Трехдиагональные матрицы и их приложения. М.: Наука. Главная редакция физико-математической литературы, 1985г. ,208 с.
6. Самарский А.А., Николаев Е.С. Методы решения сеточных уравнений. М.: Наука, 1978.
7. Фролов А.В. Ещё один метод распараллеливания прогонки с использованием ассоциативности операций // Суперкомпьютерные дни в России: Труды международной конференции (28-29 сентября 2015 г., г. Москва). – М.: Изд-во МГУ, 2015. с. 151-162
8. Фролов А.В. Использование последовательно-параллельного метода для распараллеливания алгоритмов с ассоциативными операциями // Суперкомпьютерные дни в России: Труды международной конференции (28-29 сентября 2015 г., г. Москва). – М.: Изд-во МГУ, 2015. с. 176-184
9. Фролов А.В., Антонов А.С., Воеводин Вл.В., Теплов А.М. Сопоставление разных методов решения одной задачи по методике проекта Algowiki // Параллельные вычислительные технологии (ПаВТ’2016): труды международной научной конференции (г. Архангельск, 28 марта – 1 апреля 2016 г.). Челябинск: Издательский центр ЮУрГУ, 2016. С. 347-360.

[1] в идеальных условиях концепции неограниченного параллелизма, но не в реальных условиях

10. Фролов А.В. Нециклическая редукция - незаслуженно забытый метод? // Параллельные вычислительные технологии (ПаВТ'2016): труды международной научной конференции (г. Архангельск, 28 марта – 1 апреля 2016 г.). Челябинск: Издательский центр ЮУрГУ, 2016. С. 800.
11. Stone H.S. Parallel Tridiagonal Equation Solvers // ACM Trans. on Math. Software, Vol. 1, No. 4 (Dec. 1975), P. 289-307.

Cyclic reduction graph critical path and coefficient of logarithm

A.V. Frolov[1]

INM RAS[1]

Alternative variant of cyclic reduction and its properties are the main speculations in this paper.

Keywords: critical path, algorithm graph, cyclic reduction.

References

1. Antonov A.S., Voevodin Vad.V., Voevodin Vl.V., Teplov A.M., Frolov A.V.. Pervaya versiya Otkrytoy entsiklopedii svoystv algoritmov [First Version of Algorithms' Properties' Open Encyclopedia] // Vestnik UGATU. Seriya upravlenie, vychislitel'naya tekhnika i informatika [Bulletin of UGATU. Series: Control, Compruting Technique & Informatics]. Tom 19, N 2(68), 2015., S.150-159.
2. Voevodin V.V. Vychislitel'nye osnovy lineynoy algebry [Computing Basics of Linear Algebra]. M.: Nauka, 1977.
3. Voevodin V.V., Kuznetsov Yu.A. Matritsy i vychisleniya [Matrices & Computing]. M.: Nauka, 1984.
4. Voevodin V.V. Matematicheskie osnovy parallel'nykh vychisleniy [Mathematics' Basics of Parallel Computing]. M.: Izd. Mosk. un-ta, 1991. 345 s.
5. Il'in V.P., Kuznetsov Yu.I. Trekhdiagonal'nye matritsy i ikh prilozheniya [Tridiagonal Matrices & Applications]. M.: Nauka. Glavnaya redaktsiya fiziko-matematicheskoy literatury, 1985g., 208 s.
6. Samarskiy A.A., Nikolaev E.S. Metody resheniya setochnykh uravneniy [Grid Equations Solving Methods]. M.: Nauka, 1978.
7. Frolov A.V. Eshchye odin metod rasparallelivaniya progonki s ispol'zovaniem assotsiativnosti operatsiy [Yet another Tomas Algorithm Parallelizing Method with Operathions Associativity Using]// Superkomp'yuternye dni v Rossii: Trudy mezhdunarodnoy konferentsii (28-29 sentyabrya 2015 g., g. Moskva) [Russian Supercomputer Days: Proceedings of the International Scientific Conference (Moscow, Russia, September, 28-29, 2015]. Moscow, MSU Publishing, 2015. P. 151-162
8. Frolov A.V. Ispol'zovanie posledovatel'no-parallel'nogo metoda dlya rasparallelivaniya algoritmov s assotsiativnymi operatsiyami [Sequential-Parallel Method's Using for Algorithms Containing Associative Operations Parallelizing]// Superkomp'yuternye dni v Rossii: Trudy mezhdunarodnoy konferentsii (28-29 sentyabrya 2015 g., g. Moskva) [Russian Supercomputer Days: Proceedings of the International Scientific Conference (Moscow, Russia, September, 28-29, 2015]. Moscow, MSU Publishing, 2015. P. 176-184
9. Frolov A.V., Antonov A.S., Voevodin Vl.V., Teplov A.M. Sopostavleniye raznyh metodov resheniya odnoj zadachi po metodike proekta Algowiki [One problem solving different methods' comparison according to the criteria of the Algowiki project] // Parallel'nye vychislitel'nye tekhnologii (PaVT'2016): trudy mezhdunarodnoy nauchnoy konferentsii (28 marta – 1 aprelya 2016 g., g. Arkhangelsk) [Parallel Computational Technologies (PCT'2016): Proceedings of the International Scientific Conference (Arkhangelsk, Russia, March, 28 – April, 1, 2016)]. Chelyabinsk, Publishing of the South Ural State University, 2016. P. 347–360.
10. Frolov A.V. Neciklicheskaya redukciya – nezasluzhenno zabytyy metod? [Noncyclic reduction is undeservedly forgotten method?] // Parallel'nye vychislitel'nye tekhnologii (PaVT'2016): trudy mezhdunarodnoy nauchnoy konferentsii (28 marta – 1 aprelya 2016 g., g. Arkhangelsk) [Parallel

Computational Technologies (PCT'2016): Proceedings of the International Scientific Conference (Arkhangelsk, Russia, March, 28 – April, 1, 2016)]. Chelyabinsk, Publishing of the South Ural State University, 2016. P. 800.

11. Stone H.S. Parallel Tridiagonal Equation Solvers // ACM Trans. on Math. Software, Vol. 1, No. 4 (Dec. 1975), P. 289-307.