УДК 681.322.012

Л.Д. ЕЛФИМОВА

НОВЫЙ БЫСТРЫЙ РЕКУРСИВНЫЙ АЛГОРИТМ УМНОЖЕНИЯ МАТРИЦ

Аннотация. Предложен новый рекурсивный алгоритм умножения матриц порядка $n = 2^q$ $(q > 1)$, в котором в качестве базового применяется быстрый гибридный алгоритм умножения матриц порядка 4μ при $\mu = 2^{q-1}$ $(q > 0)$. По сравнению с известными рекурсивными алгоритмами Штрассена и Винограда–Штрассена данный алгоритм позволяет минимизировать на 7% мультипликативную сложность, равную $W_{\text{м}} \approx 0.932 n^{2.807}$ операций умножения на глубине рекурсии $d = \log_2 n - 3$, и сократить вектор вычислений на три рекурсивных шага. Дана оценка мультипликативной сложности представленного алгоритма.

Ключевые слова: линейная алгебра, блочно-рекурсивные алгоритмы Штрассена и Винограда–Штрассена, семейство быстрых гибридных алгоритмов умножения матриц.

В настоящее время при возрастающих требованиях к скорости решения современных задач, где необходимы большие объемы вычислений с плотными матрицами, актуальна проблема получения быстрых эффективных матричных алгоритмов, в которых мультипликативная сложность является доминирующей.

Среди известных быстрых алгоритмов умножения матриц на практике эффективны регулярный алгоритм Винограда [1] и блочно-рекурсивные алгоритмы Штрассена [2], Винограда–Штрассена [3] и Лейдермана [4], которые используются во многих предметных областях, особенно в цифровой обработке сигналов и изображений для высокоскоростных вычислений в реальном времени.

Первый быстрый регулярный алгоритм умножения $(n \times n)$-матриц разработал Виноград [1], мультипликативная сложность которого составляет $O(0.5n^3 + n^2)$ операций умножения, при этом аддитивная и общая сложности соответственно равны $O(1.5n^3 + 2n^2 - 2n)$ операций сложения и $O(2n^3 + 3n^2 - 2n)$ операций умножения/сложения. В 1969 г. Штрассен предложил быстрый блочно-рекурсивный алгоритм умножения матриц порядка $n = 2^q$ $(q > 0)$ [2], мультипликативная сложность которого равна $O(n^{\log_2 7}) \sim O(n^{2.807})$ операций умножения. В качестве базового алгоритма Штрассен использовал алгоритм умножения (2×2)-матриц, требующий семь операций умножения и 18 операций сложения. Данный рекурсивный алгоритм реализуется за $\log_2 n$ рекурсивных шагов и эффективен при большом порядке матриц $(n \geq 64)$. В 1971 г. Виноград предложил модифицированную версию [3] алгоритма Штрасссена, которая позволила минимизировать аддитивную сложность алгоритма [2] при неизменной мультипликативной сложности $O(n^{2.807})$ операций умножения. В 1976 г. Лейдерман разработал блочно-рекурсивный алгоритм умножения матриц порядка $n = 3^q$ $(q > 0)$ [4], мультипликативная сложность которого составляет $O(n^{\log_3 23}) \sim O(n^{2.854})$ операций умножения. В качестве базового алгоритма Лейдерман применил алгоритм умножения (3×3)-матриц, требующий 23 операции умножения и 98 операции сложения. Данный рекурсивный алгоритм реализуется за $\log_3 n$ рекурсивных шагов и применяется для матриц нечетного порядка. Кроме того, в настоящее время разработано семейство новых быстрых алгоритмов умножения матриц порядка $n = 2\mu$ $(\mu > 1)$ [5, 6], $n = 3\mu$ $(\mu > 1)$ [7], $n = 4\mu$ $(\mu > 0)$ [5, 6], $n = 6\mu$ $(\mu > 0)$ [7], при построении которых использовались алгоритмы умножения (2×2)-матриц Штрассена [2] и Винограда–Штрассена [3], (3×3)-матриц Лейдермана [4] и алгоритм для скалярного произведения Винограда [1].

Данные гибридные алгоритмы, полученные новым способом построения быстрых алгоритмов, сочетают преимущества перечисленных известных алгоритмов. Однако их отличительной особенностью является то, что они характеризуются наименьшей операционной сложностью, обусловленной впервые достигнутой в них одновременной минимизацией мультипликативной и аддитивной сложностей. Быстрый гибридный алгоритм умножения матриц порядка $n = 4\mu$ $(\mu > 0)$ [6] по сравнению с алгоритмом Винограда [1] отличается минимизированными, мультипликативной, аддитивной и общей сложностями, которые соответственно составляют $W_{\text{м}} = 0.4375n^3 + 1.75n^2$ операций умножения, $W_{\text{а}} = 1.3125n^3 + 7.25n^2 - 7n$ операций сложения, $W_{\text{общ}} = 1.75n^3 + 9n^2 - 7n$ операций умножения/сложения. По сравнению с рекурсивным алгоритмом Штрассена [2] данный гибридный алгоритм характеризуется минимизированной мультипликативной сложностью при $n = 8, 16, 32$.

Цель настоящей статьи — оптимизация мультипликативной сложности известных алгоритмов умножения матриц. Рассмотрен новый быстрый рекурсивный алгоритм умножения матриц порядка $n = 2^q$ $(q > 1)$, который построен на основе быстрого гибридного алгоритма порядка $n = 4\mu$ при $\mu = 2^{q-1}$ $(q > 0)$ и позволяет получить минимизированную на 7% по сравнению с блочно-рекурсивными алгоритмами Штрассена и Винограда–Штрассена мультипликативную сложность, равную $W_{\text{м}} \approx 0.932n^{2.807}$ операций умножения на глубине рекурсии $d = \log_2 n - 3$, при этом вектор вычислений сокращается на три рекурсивных шага. Дана оценка мультипликативной сложности предложенного алгоритма.

Быстрый рекурсивный алгоритм умножения матриц порядка $n = 2^q$ $(q > 1)$. При построении рекурсивного алгоритма умножения квадратных $(n\times n)$-матриц $A = (a_{ij})$ и $B = (b_{ij})$ порядка $n = 2^q$ $(q > 1)$ в качестве базового используется быстрый гибридный алгоритм умножения матриц порядка $n = 4\mu$ при $\mu = 2^{q-1}$ $(q > 0)$ [6].

Рассмотрим первый шаг рекурсии данного алгоритма. Согласно базовому алгоритму вначале вычисляем матрицы $S^1 = (s_{ij}^1), \ldots, S^8 = (s_{ij}^8)$ порядка $n/2$. Элементами данных матриц являются коэффициенты $s_{ik}^1, \ldots, s_{ik}^4$ и $s_{kj}^5, \ldots, s_{kj}^8$, определяемые по следующим формулам:

$$\begin{aligned}
&s_{ik}^1 = a_{2i,2k-1} + a_{2i,2k}, && s_{kj}^5 = b_{2k-1,2j} - b_{2k-1,2j-1},\\
&s_{ik}^2 = s_{ik}^1 - a_{2i-1,2k-1}, && s_{kj}^6 = b_{2k,2j} - s_{kj}^5,\\
&s_{ik}^3 = a_{2i-1,2k-1} - a_{2i,2k-1}, && s_{kj}^7 = b_{2k,2j} - b_{2k-1,2j},\\
&s_{ik}^4 = a_{2i-1,2k} - s_{ik}^2, && s_{kj}^8 = s_{kj}^6 - b_{2k,2j-1},\quad i,j,k = 1,2,\ldots,n/2.
\end{aligned} \tag{1}$$

Затем вычисляем семь произведений $P^1 = (p_{ij}^1), \ldots, P^7 = (p_{ij}^7)$ двух матриц порядка $n/2$:

$$
\begin{aligned}
p_{ij}^1 &= z_{ij}^1 - \varphi_i^1 - \omega_j^1, \\
p_{ij}^2 &= z_{ij}^2 - \varphi_i^2 - \omega_j^2, \\
p_{ij}^3 &= z_{ij}^3 - \varphi_i^3 - \omega_j^3, \\
p_{ij}^4 &= z_{ij}^4 - \varphi_i^4 - \omega_j^4, \\
p_{ij}^5 &= z_{ij}^5 - \varphi_i^5 - \omega_j^5, \\
p_{ij}^6 &= z_{ij}^6 - \varphi_i^6 - \omega_j^6, \\
p_{ij}^7 &= z_{ij}^7 - \varphi_i^7 - \omega_j^7, \quad i, j = 1, 2, \ldots, n/2,
\end{aligned}
\tag{2}
$$

где

$$
\begin{aligned}
z_{ij}^1 &= \sum_{k=1}^{n/4} (s_{i,2k-1}^2 + s_{2k,j}^6)(s_{2k-1,j}^6 + s_{i,2k}^2), \\
z_{ij}^2 &= \sum_{k=1}^{n/4} (a_{2i-1,4k-3} + b_{4k-1,2j-1})(b_{4k-3,2j-1} + a_{2i-1,4k-1}), \\
z_{ij}^3 &= \sum_{k=1}^{n/4} (a_{2i-1,4k-2} + b_{4k,2j-1})(b_{4k-2,2j-1} + a_{2i-1,4k}), \\
z_{ij}^4 &= \sum_{k=1}^{n/4} (s_{i,2k-1}^3 + s_{2k,j}^7)(s_{2k-1,j}^7 + s_{i,2k}^3), \\
z_{ij}^5 &= \sum_{k=1}^{n/4} (s_{i,2k-1}^1 + s_{2k,j}^5)(s_{2k-1,j}^5 + s_{i,2k}^1), \\
z_{ij}^6 &= \sum_{k=1}^{n/4} (s_{i,2k-1}^4 + b_{4k,2j})(b_{4k-2,2j} + s_{i,2k}^4), \\
z_{ij}^7 &= \sum_{k=1}^{n/4} (a_{2i,4k-2} + s_{2k,j}^8)(s_{2k-1,j}^8 + a_{2i,4k}), \quad i, j = 1, 2, \ldots, n/2, \quad k = 1, 2, \ldots, n/4;
\end{aligned}
\tag{3}
$$

$$
\varphi_i^1 = \sum_{k=1}^{n/4} s_{i,2k-1}^2 \cdot s_{i,2k}^2, \qquad \omega_j^1 = \sum_{k=1}^{n/4} s_{2k,j}^6 \cdot s_{2k-1,j}^6,
$$

$$
\varphi_i^2 = \sum_{k=1}^{n/4} a_{2i-1,4k-3} a_{2i-1,4k-1}, \quad \omega_j^2 = \sum_{k=1}^{n/4} b_{4k-1,2j-1} b_{4k-3,2j-1},
$$

$$
\varphi_i^3 = \sum_{k=1}^{n/4} a_{2i-1,4k-2} a_{2i-1,4k}, \quad \omega_j^3 = \sum_{k=1}^{n/4} b_{4k,2j-1} b_{4k-2,2j-1},
$$

$$\varphi_i^4 = \sum_{k=1}^{n/4} s_{i,2k-1}^3 \cdot s_{i,2k}^3, \qquad \omega_j^4 = \sum_{k=1}^{n/4} s_{2k,j}^7 \cdot s_{2k-1,j}^7, \tag{4}$$

$$\varphi_i^5 = \sum_{k=1}^{n/4} s_{i,2k-1}^1 \cdot s_{i,2k}^1, \qquad \omega_j^5 = \sum_{k=1}^{n/4} s_{2k,j}^5 \cdot s_{2k-1,j}^5,$$

$$\varphi_i^6 = \sum_{k=1}^{n/4} s_{i,2k-1}^4 \cdot s_{i,2k}^4, \qquad \omega_j^6 = \sum_{k=1}^{n/4} b_{4k,2j} b_{4k-2,2j},$$

$$\varphi_i^7 = \sum_{k=1}^{n/4} a_{2i,4k-2} a_{2i,4k}, \qquad \omega_j^7 = \sum_{k=1}^{n/4} s_{2k,j}^8 \cdot s_{2k-1,j}^8,$$

$$i, j = 1, 2, \ldots, n/2, \; k = 1, 2, \ldots, n/4.$$

Для нахождения результирующей матрицы C вначале выполняем следующие промежуточные вычисления:

$$t_{ij}^1 = p_{ij}^1 + p_{ij}^2, \; t_{ij}^2 = t_{ij}^1 + p_{ij}^4, \; t_{ij}^3 = p_{ij}^5 + p_{ij}^6, \; i, j = 1, 2, \ldots, n/2. \tag{5}$$

Затем вычисляем элементы результирующей матрицы $C = (c_{ij})$ по следующим формулам:

$$c_{2i-1,2j-1} = p_{ij}^2 + p_{ij}^3,$$

$$c_{2i-1,2j} = t_{ij}^1 + t_{ij}^3, \tag{6}$$

$$c_{2i,2j-1} = t_{ij}^2 - p_{ij}^7,$$

$$c_{2i,2j} = t_{ij}^2 + p_{ij}^5 \;, \; i, j = 1, 2, \ldots, n/2.$$

Базовый алгоритм (1)–(6) можно также использовать в рекурсивном режиме для вычисления промежуточных матриц $P^1, \ldots, P^7$. На втором шаге рекурсии алгоритм оперирует матрицами порядка $n/4$, вычисляя матрицы коэффициентов по формуле (1) и формируя промежуточные матрицы $P^1 \sim P^7$.

На каждом последующем шаге порядок оперируемых матриц уменьшается вдвое. На последнем шаге рекурсии алгоритм оперирует (4×4)-матрицами, выполняя вычисления согласно алгоритму (1)–(6). При реализации полного рекурсивного вычислительного процесса алгоритм требует $(\log_2 n - 1)$ рекурсивных шагов.

Оценка мультипликативной сложности. Важнейшей характеристикой, влияющей на оптимизацию вычислительной сложности рекурсивного алгоритма, является глубина рекурсии d. Оценим мультипликативную сложность нового алгоритма с учетом данной величины. Поскольку базовый алгоритм (1)–(6) имеет мультипликативную сложность, равную $W_{\text{м}}^{\text{баз}} = 0.4375n^3 + 1.75n^2$ операций умножения, мультипликативную сложность построенного на его основе рекурсивного алгоритма определяем по следующей формуле:

$$W_{\text{м}}^{\text{рек}} = 7^{(\log_2 n - \omega)} \cdot [0.4375(2^\omega)^3 + 1.75(2^\omega)^2] = 7^{\log_2 n} \cdot \left[\frac{0.4375(2^\omega)^3 + 1.75(2^\omega)^2}{7^\omega} \right] =$$

$$= \alpha \cdot n^{\log_2 7} \sim \alpha \cdot n^{2.807} \text{ операций умножения,} \tag{7}$$

где

$$\alpha = \frac{0.4375(2^{\omega})^3 + 1.75(2^{\omega})^2}{7^{\omega}} \text{ — коэффициент при } n^{2.807},\ \omega = 2, 3, \ldots, \log_2 n.$$

При этом глубина рекурсии составляет $d = (\log_2 n - \omega - 1)$ рекурсивных шагов.

Максимальный процент минимизации мультипликативной сложности достигается при $\omega = 4$. Используя формулу (7), определяем мультипликативную сложность нового алгоритма при $\omega = 4$ на глубине рекурсии $d = \log_2 n - 3$:

$$W_{\text{м}}^{\text{рек}} = 7^{(\log_2 n - 4)} \cdot [0.4375(2^4)^3 + 1.75(2^4)^2] = \frac{2240}{7^4} \cdot 7^{\log_2 n} =$$

$$= \frac{2240}{2401} n^{\log_2 7} \approx 0.932 n^{2.807} \text{ операций умножения.}$$

Сокращая вектор вычислений еще на один шаг, получаем $\alpha \approx 0.959$ при $\omega = 5$; $d = \log_2 n - 4$. При увеличении вектора вычислений коэффициент α увеличивается, а именно $\alpha \approx 0.979$ при $\omega = 3$; $d = \log_2 n - 2$.

При реализации полного рекурсивного процесса, в котором алгоритм оперирует (4×4)-матрицами, характеристики алгоритма при $\omega = 2$ составляют $\alpha \approx 1.143$; $d = \log_2 n - 1$.

Таким образом, в работе построен новый рекурсивный алгоритм умножения матриц порядка $n = 2^q$ $(q > 1)$, который использует в качестве базового быстрый гибридный алгоритм умножения матриц порядка $n = 4\mu$ при $\mu = 2^{q-1}$ $(q > 0)$. Новый подход к построению рекурсивных алгоритмов дает возможность получить наименьшую мультипликативную сложность $O(0.932 n^{2.807})$ по сравнению с блочно-рекурсивными алгоритмами Штрассена и Винограда–Штрассена, мультипликативная сложность которых равна $O(n^{2.807})$.

Представленный алгоритм обеспечивает максимальный процент минимизации ($\approx$ 7%) мультипликативной сложности на глубине рекурсии, равной $d = \log_2 n - 3$, сокращая вектор вычислений на три рекурсивных шага.

Построенный на основе гибридного алгоритма новый рекурсивный алгоритм расширяет семейство быстрых гибридных алгоритмов, имеет практическую ценность, его можно эффективно реализовать на различных вычислительных системах.

СПИСОК ЛИТЕРАТУРЫ

1. Winograd S.A. A new algorithm for inner product. *IEEE Transactions on Computers*. 1968. Vol. C-18. P. 693–694.
2. Strassen V. Gaussian elimination is not optimal. *Numerische Mathematik*. 1969. Vol. 13. P. 354–356.
3. Winograd S. On multiplication of 2×2 matrices. *Linear Algebra and Its Application*. 1971. Vol. 4. P. 381–388.
4. Laderman J.D. A noncommutative algorithm for multiplying 3×3 matrices using 23 multiplications. *Bulletin of the American Mathematical Society*. 1976. Vol. 82, N 1. P. 126–128.
5. Елфимова Л.Д., Капитонова Ю.В. Быстрый алгоритм для умножения матриц и его эффективная реализация на систолических массивах. *Кибернетика и системный анализ*. 2001. № 1. С. 135–150.
6. Елфимова Л.Д. Быстрые гибридные алгоритмы умножения матриц. *Кибернетика и системный анализ*. 2010. Т. 46, № 4. С. 49–59.
7. Елфимова Л.Д. Новые быстрые гибридные алгоритмы умножения матриц. *Кибернетика и системный анализ*. 2011. Т. 47, № 6. С. 59–67.

Надійшла до редакції 22.10.2018

Л.Д. Єлфімова

НОВИЙ ШВИДКИЙ РЕКУРСИВНИЙ АЛГОРИТМ МНОЖЕННЯ МАТРИЦЬ

Анотація. Запропоновано новий рекурсивний алгоритм множення матриць порядку $n = 2^q$ $(q > 1)$, в якому як базовий застосовується швидкий гібридний алгоритм множення матриць порядку $n = 4\mu$, коли $\mu = 2^{q-1}$ $(q > 0)$. Порівняно з відомими рекурсивними алгоритмами Штрасена та Винограда–Штрасена цей алгоритм дозволяє мінімізувати на 7% мультиплікативну складність, яка дорівнює $W_{\text{м}} \approx 0.932n^{2.807}$ операцій множення на глибині рекурсії $d = \log_2 n - 3$, та скоротити вектор обчислень на три рекурсивних кроки. Наведено оцінку мультиплікативної складності представленого алгоритму.

Ключові слова: лінійна алгебра, блочно-рекурсивні алгоритми Штрасена та Винограда–Штрасена, сім'я швидких гібридних алгоритмів множення матриць.

L.D. Jelfimova

A NEW FAST RECURSIVE MATRIX MULTIPLICATION ALGORITHM

Abstract. A new recursive algorithm is proposed for multiplying matrices of order $n = 2^q$ $(q > 1)$. This algorithm is based on fast hybrid algorithm for multiplying matrices of order $n = 4\mu$ for $\mu = 2^{q-1}$ $(q > 0)$. As compared with the well-known recursive Strassen's and Winograd–Strassen's algorithms, the new algorithm minimizes by 7% the multiplicative complexity equal to $W_{\text{м}} \approx 0.932n^{2.807}$ multiplication operations at recursive level $d = \log_2 n - 3$ and reduces the computation vector by three recursive steps. The multiplicative complexity of the algorithm is estimated.

Keywords: linear algebra, Strassen's and Winograd–Strassen's block-recursive matrix multiplication algorithms, family of fast hybrid matrix multiplication algorithms.

Елфимова Лариса Дмитриевна,
младший научный сотрудник Института кибернетики им. В.М. Глушкова НАН Украины, Киев, e-mail: larisaelf@gmail.com.