УДК 510.52

О ВОЗМОЖНОМ МАТРИЧНОМ ПРЕДСТАВЛЕНИИ АНАЛИТИЧЕСКОГО РЕШЕНИЯ ОДНОГО НЕЛИНЕЙНОГО РЕКУРРЕНТНОГО СООТНОШЕНИЯ

ABOUT A POSSIBLE MATRIX REPRESENTATION OF ANALYTICAL SOLUTIONS OF A NONLINEAR RECURRENT RELATION

Головешкин В.А., д-р. техн. наук, проф.,

Пономарёв А.В., кандидат физ.-мат. наук, доц.,

Московский государственный университет приборостроения и информатики

Ульянов М.В., д-р. техн. наук, проф.

Московский государственный университет печати имени Ивана Федорова

Аннотация

В статье рассматривается аналитическое решение одного нелинейного рекуррентного соотношения с квадратичной аддитивной функцией. В целом такие рекуррентные соотношения характерны для функций трудоемкости рекурсивных алгоритмов, разработанных методом декомпозиции. Квадратичная аддитивная функция возникает в рекурсивных алгоритмах, реализующих матричные операции, в частности в известном алгоритме умножения матриц по Штрассену. Для такого нелинейного рекуррентного соотношения предложено матричное представление аналитического решения, которое имеет определенный теоретический интерес.

Abstract

The article discusses the analytical solution of a nonlinear recurrent equation with a quadratic additive function. In general, the recurrence relations are characteristic for the complexity of recursive algorithms, developed by the methods of decompositions. The quadratic additive function occurs in recursive algorithms, implemented using matrix operations, in particular in a well-known algorithm for multiplication of matrices of Strassen. For such a non-linear recurrent equation the proposed matrix representation of the analytical solution, which has the prominent theoretical interest.

1. Введение

Ресурсный анализ рекурсивных алгоритмов предполагает получение аналитических решений для функций трудоемкости алгоритмов или оценок этих функций на основе теоретического анализа. Практический интерес к этой задаче порождается необходимостью исследования класса рекурсивных алгоритмов, разработанных методом декомпозиции. Основная идея метода декомпозиции состоит в сведении исходной задачи к решению ряда более простых задач с понижением их размерности [1, 2] и последующим объединением полученных решений, что приводит к алгоритмам, имеющим рекурсивную структуру и описанию функции трудоёмкости в виде рекуррентного соотношения [1].

Формально: пусть n — размерность решаемой задачи, функция $d(n)$ описывает трудоемкость фрагмента алгоритма, выполняющего декомпозицию задачи, а функция $U(n)$ — трудоемкость фрагмента алгоритма, объединяющего полученные решения. При этом на шаге останова рекурсии для некоторой малой размерности задачи, т. е. при $0 \le n \le n_0$, возможно ее прямое (не рекурсивное) решение с трудоемкостью $h(n), 0 \le n \le n_0$, а собственно рекурсия состоит в таком разделении задачи, которое приводит к необходимости решения C подзадач размерностью n/k. Размерность решаемой задачи должна, очевидно, быть целой, следовательно вместо значения n/k в качестве аргумента функции должна фигурировать целая часть частного с округлением вниз или вверх, т. е. $\lfloor n/k \rfloor$ или $\lceil n/k \rceil$, что существенно осложняет получение аналитического решения. Вводя обозначение $g(n) = d(n) + U(n)$ приходим к следующему рекуррентному соотношению на трудоемкость

$$\begin{cases} f_A(n) = h(n) \ , \quad 0 \le n \le n_0; \\ f_A(n) = C \cdot f_A([n/k]) + g(n) \, , n > n_0, \end{cases} \tag{1}$$

где через $[\cdot]$ мы обозначаем $\lfloor \cdot \rfloor$ или $\lceil \cdot \rceil$. Функция $g(n)$ в (1) называется аддитивной функцией рекуррентного соотношения.

Оценка асимптотического поведения функции трудоемкости, заданного соотношением (1) может быть получена на основе известной теоремы о рекуррентных соотношениях (J.L. Bentley, D. Haken, J.B. Saxe, 1980 [3]). Для функций $g(n)$ определённого вида могут быть получены не только асимптотические оценки, но и точные аналитические решения. Ряд результатов в этом направлении получен авторами статьи в [4,5], в частности в [5] были получены аналитические решения для функции $g(n)$ степенного вида. Рассмотрение частного случая (1) — случая с квадратичной аддитивной функцией показало, что аналитическое решение может быть представлено в матричной форме. Изложению этого результата и посвящена настоящая статья.

2. Постановка задачи

Рассмотрим нелинейное рекуррентное соотношение вида (1) с квадратичной аддитивной функцией $g(n) = A \cdot n^2$. Поскольку множество полиномов замкнуто относительно сложения и умножения, то, не теряя общности, мы можем рассматривать $g(n) = n^2$. В дальнейшем функцию трудоёмкости алгоритма будем обозначать через f. Рассмотрим далее размерности вида $\lfloor n/k \rfloor$ при значениях $C \ge 2$, $k \ge 2, n_0 = 0, h(0) = \alpha$. В этой постановке аналитическое решение получено авторами в [5], а задача настоящей статьи состоит в том, чтобы показать, что для аналитического решения рекуррентного соотношения

$$\begin{cases} f(0) = \alpha, n = 0; \\ f(n) = C \cdot f(\lfloor n/k \rfloor) + n^2, n > 0, \end{cases} \tag{2}$$

возможно получение специального матричного представления, которое, по мнению авторов, имеет определенный теоретический интерес.

Отметим, что соотношения типа (2) характерны для функций трудоемкости алгоритмов, реализующих матричные операции, например, — для алгоритма умножения матриц, предложенного В. Штрассеном в 1969 г. [6]. Рекуррентное соотношение, описывающее трудоёмкость этого алгоритма (при $n = 2^k$) имеет следующий вид:

$$f(n) = 7 \cdot f(n/2) + A \cdot n^2 .$$

3. Матричное представление аналитического решения

Первоначально будем использовать подход, предложенный нами в [4,5], а именно введем в рассмотрение специальное представление числа n по степеням k:

$$n = b_m k^m + b_{m-1} k^{m-1} + \cdots + b_1 k + b_0 = \sum_{i=0}^{m} b_i k^i , \quad m = \lfloor \log_k n \rfloor, m \ge 1 , \tag{3}$$

где b_m: $0 < b_m < k$, b_i: $0 \le b_i < k, i = \overline{0, m-1}$ — цифры в системе счисления по основанию k.

Подставим полученное представление (3) в рекуррентное соотношение (2), а именно в основное соотношение: $f(n) = C \cdot f(\lfloor n/k \rfloor) + n^2$. Имеем:

$$\begin{aligned} f\left(\sum_{i=0}^{m} b_i k^i\right) &= C \cdot f\left(\sum_{i=1}^{m} b_i k^{i-1}\right) + \left(\sum_{i=0}^{m} b_i k^i\right)^2 = \\ &= C^2 \cdot f\left(\sum_{i=2}^{m} b_i k^{i-2}\right) + C \cdot \left(\sum_{i=1}^{m} b_i k^{i-1}\right)^2 + \left(\sum_{i=0}^{m} b_i k^i\right)^2 = \\ &= C^3 \cdot f\left(\sum_{i=3}^{m} b_i k^{i-3}\right) + C^2 \cdot \left(\sum_{i=2}^{m} b_i k^{i-2}\right)^2 + C \cdot \left(\sum_{i=1}^{m} b_i k^{i-1}\right)^2 + \left(\sum_{i=0}^{m} b_i k^i\right)^2 = \\ &= \sum_{j=0}^{m} C^j \left(\sum_{i=j}^{m} \left(b_i k^{i-j}\right)^2\right) + C^{m+1} \cdot \alpha \end{aligned} \tag{4}$$

Заметим, что основной компонент решения — выражение

$$S(b_0, b_1, \ldots, b_m) = \sum_{j=0}^{m} C^j \left(\sum_{i=j}^{m} b_i k^{i-j}\right)^2 \tag{5}$$

является квадратичной формой относительно элементов $b_i, 0 \le i \le m, b_m \ne 0$ — цифр представления числа n по степеням k в соответствии с (3).

Введём в рассмотрение представление некоторой квадратичной формы $(a_1 x_1 + a_2 x_2 + \cdots + a_n x_n)^2$ в виде произведения векторов. Обозначим через $\mathbf{A}$ вектор столбец коэффициентов, тогда вектор строка $\mathbf{A^T}$ имеет вид: $(a_1, a_2, \cdots, a_n)$. Аналогично введем вектор столбец неизвестных $\mathbf{X}$ и его транспонированное представление — вектор строку $\mathbf{X^T}$: $(x_1, x_2, \cdots, x_n)$. В этих обозначениях

$$(a_1 x_1 + a_2 x_2 + \cdots + a_n x_n) = \mathbf{A^T X} = \mathbf{X^T A} ,$$

следовательно квадратичная форма представима в виде

$$(a_1x_1 + a_2x_2 + \dots + a_nx_n)^2 = \mathbf{X^T A A^T X} = \mathbf{X^T D X},$$

при этом $\mathbf{D} = \mathbf{AA^T}$ есть матрица, как результат умножения столбца на строку, ее размерность равна $n \times n$, а элементы этой матрицы имеют вид: $d_{ij} = a_i \cdot a_j$.

Для матричного представления квадратичной формы (5) введем специальный вектор столбец $\mathbf{A_j}$ размерности $m+1$, $\mathbf{A_j}^\mathbf{T} = (a_{0j}, a_{1j}, \dots, a_{mj})$, компоненты a_{ij} которого определены следующим образом:

$$a_{ij} = \begin{cases} 0, & i < j, \\ k^{i-j}, & j \le i \le m. \end{cases} \tag{6}$$

Матрицу, получаемую при произведении $\mathbf{A_j} \cdot \mathbf{A_j}^\mathbf{T}$ обозначим через $\mathbf{D_j}$, ее размерность равна $(m+1)\times(m+1)$ и введем вектор столбец $\mathbf{B} : \mathbf{B^T} = (b_0, b_1, \dots, b_m)$. В указанных обозначениях мы получаем следующее матричное представление для основного компонента аналитического решения рекуррентного соотношения (2)

$$\begin{aligned} S(b_0, b_1, \dots, b_m) &= \sum_{j=0}^{m} C^j \left(\sum_{i=j}^{m} b_i k^{i-j} \right)^2 = \\ &= \sum_{j=0}^{m} C^j \left(\mathbf{B^T D_j B} \right) = \mathbf{B^T} \left(\sum_{j=0}^{m} C^j \mathbf{D_j} \right) \mathbf{B} = \\ &= \mathbf{B^T} \cdot \mathbf{D} \cdot \mathbf{B}. \end{aligned} \tag{7}$$

Отметим, что в (7) матрица $\mathbf{D}$ есть сумма матриц $\mathbf{D_j}$ с коэффициентами C^j:

$$\mathbf{D} = \sum_{j=0}^{m} \left(C^j \mathbf{D_j} \right). \tag{8}$$

4. Вид элементов матрицы D

Получим явные выражения для элементов матрицы $\mathbf{D}$, отметив при этом, что общий вид диагональных и недиагональных элементов различен, и указав, что в силу (6), вектор столбец $\mathbf{A_j}$ размерности $m+1$, образующий матрицу $\mathbf{D_j} = \mathbf{A_j} \cdot \mathbf{A_j}^\mathbf{T}$, имеет нулевые элементы.

Для диагональных элементов d_{ss} матрицы $\mathbf{D}$ в силу (6) и (8) получаем:

$$d_{ss} = \sum_{i=0}^{m} C^i (a_{si} \cdot a_{si}) = \sum_{i=0}^{s} C^i (a_{si} \cdot a_{si}) = \sum_{i=0}^{s} C^i \left(k^{2s-2i} \right) = k^{2s} \sum_{i=0}^{s} \frac{C^i}{k^{2i}}. \tag{9}$$

при этом в (9) возникают два случая, связанные с соотношением между C и k^2, которые мы рассмотрим отдельно:

1. Случай $k^2 \neq C$. В этих условия мы можем вычислить последнюю сумму в (9) как частичную сумму геометрической прогрессии:

$$d_{ss} = k^{2s} \sum_{i=0}^{s} \frac{C^i}{k^{2i}} = k^{2s} \frac{1-\left(\frac{C}{k^2}\right)^{s+1}}{1-\frac{C}{k^2}} = \frac{k^{2(s+1)} - C^{(s+1)}}{k^2 - C}.$$

2. Случай $k^2 = C$. При этом отношение внутри суммы в (9) равно 1, а сама сумма равна $s+1$:

$$d_{ss} = k^{2s} \sum_{i=0}^{s} \frac{C^i}{k^{2i}} = (s+1) \cdot k^{2s}.$$

Для элементов d_{st} матрицы $\mathbf{D}$, лежащих ниже главной диагонали, т.е. при $s > t$ в силу (6) и (8) имеем:

$$d_{st} = \sum_{i=0}^{m} C^i (a_{si} \cdot a_{ti}) = \sum_{i=0}^{t} C^i (a_{si} \cdot a_{ti}) = \sum_{i=0}^{t} C^i \left(k^{s-i} \cdot k^{t-i}\right) = k^{s+t} \sum_{i=0}^{t} \frac{C^i}{k^{2i}}. \tag{10}$$

при этом в (10) аналогично d_{ss} возникают два случая, а именно:

1. Случай $k^2 \neq C$. Это неравенство позволяет вычислить последнюю сумму в (10) как частичную сумму геометрической прогрессии:

$$d_{st} = k^{s+t} \sum_{i=0}^{t} \frac{C^i}{k^{2i}} = k^{s+t} \frac{1-\left(\frac{C}{k^2}\right)^{t+1}}{1-\frac{C}{k^2}} = k^{s-t} \frac{k^{2(t+1)} - C^{(t+1)}}{k^2 - C}.$$

2. Случай $k^2 = C$, тогда отношение внутри суммы в (10) равно 1, а сама сумма равна $t+1$:

$$d_{st} = k^{s+t} \sum_{i=0}^{t} \frac{C^i}{k^{2i}} = (t+1) \cdot k^{s+t}.$$

4. Частный случай размерности: целая степень

Определенный интерес представляет особая ситуация в предложенном матричном представлении. Эта ситуация связана с тем, что в программных реализациях рекурсивных алгоритмов принято дополнять задачу до размерности n, равной целой степени k:

$$\hat{n} = k^{\lceil \log_k n \rceil}. \tag{11}$$

Такой подход позволяет избежать обработки особых ситуаций «пола» и «потолка» при организации рекурсии, и оказывается в ряде случаев более эффективным, чем особая обработка ситуаций с подзадачами четной и нечетной размерности. Это приводит вместо представления трудоемкости в виде $f(n) = C \cdot f([n/k]) + n^2$ к виду $f(n) = C \cdot f(n/k) + n^2$, причем в силу (11) результат деления n/k на каждом шаге рекурсии является целым числом.

Например, при реализации алгоритма Штрассена [1] (для которого $k = 2$) перемножаемые матрицы размерностью $n \times n$ предварительно дополняются нулями до размерности $\hat{n} \times \hat{n}$, где $\hat{n} = k^{\lceil \log_2 n \rceil}$ после чего рекурсия выполняется без анализа четности/нечетности текущей длины входа.

В рамках рассматриваемого матричного представления этот подход позволяет явно указать вид вектора столбца $\mathbf{B}$. Действительно из представления n в виде (3)

$$n = b_m k^m + b_{m-1} k^{m-1} + \cdots + b_1 k + b_0 = \sum_{i=0}^{m} b_i k^i ,$$

где $m = \lfloor \log_k n \rfloor$, и (11) в общем случае, когда само исходное значение n не есть целая степень k, т.е. при $n > k^m$ следует, что $\hat{n} = 1 \cdot k^{m+1} + 0 \cdot k^m + \cdots + 0 \cdot k + 0 = k^{m+1}$, следовательно

$$\mathbf{B}^{\mathbf{T}} = (b_0 = 0, b_1 = 0, \ldots, b_m = 0, b_{m+1} = 1) = (0, \ldots, 0, 1) \ \left|\mathbf{B}^{\mathbf{T}}\right| = m + 2 .$$

Подставляя полученное представление $\mathbf{B}$ в (7) с учетом (4) получаем аналитическое решение для случая $k^2 \neq C$ в виде:

$$f(n) = C^{m+2} \cdot \alpha + \frac{k^{2(m+2)} - C^{(m+2)}}{k^2 - C} ,$$

и для случая $k^2 = C$:

$$f(n) = C^{m+2} \cdot \alpha + (m+2) \cdot k^{2(m+1)} ,$$

при этом мы считаем, что $m = \log_k \hat{n} - 1$.

5. Заключение

В статье рассмотрено аналитическое решение нелинейного рекуррентного соотношения для функции трудоемкости рекурсивного алгоритма, построенного методом декомпозиции с квадратичной аддитивной функцией. Показано, что в этом частном случае возможно матричное представление аналитического решения и получены явные выражения для элементов соответствующей матрицы. Дополнительно рассмотрено аналитическое решение при реализации рекурсивного алгоритма с длиной входа равной целой степени k. Интересно, что формулы для значений элементов матрицы аналитического решения имею различный вид в зависимости от выполнения сравнения $k^2 = C$. На основе полученной матрицы при фиксированных в конкретном алгоритме значениях k, C можно проводить исследование зависимости функции трудоемкости от цифр представления длины входа алгоритма в системе счисления по основанию k, поскольку элементы аналитической матрицы показывают степень влияния этих цифр в соответствующих позициях на трудоемкость алгоритма.

Библиографический список

1. Кормен Т., Лейзерсон Ч., Ривест Р., Штайн К. Алгоритмы: построение и анализ, 2-ое издание: Пер. с англ. — М.: Издательский дом «Вильямс», 2005. — 1296 с.

2. Головешкин В. А., Ульянов М.В. Теория рекурсии для программистов. — М.: ФИЗМАТЛИТ. 2006. 296 с.

3. J.L. Bentley, D. Haken, J. Saxe A general method for solving divide-anal-conquer recurrences // SIGACT News, 12(3), pp. 36–44, 1980.

4. Головешкин В.А., Ульянов М.В. Аналитическое решение специального класса рекуррентных соотношений в целях анализа рекурсивных алгоритмов // Вестник Самарского государственного университета. Естественнонаучная серия. 2008. № 3(62). С. 96–107.

5. Головешкин В.А., Пономарёв А.В., Ульянов М.В., Аналитическое решение класса рекуррентных соотношений с аддитивной функцией степенного вида в целях анализа трудоёмкости рекурсивных алгоритмов // Автоматизация и современные технологии. 2011. № 3. С. 25–29.

6. Strassen V. Gaussian Elimination is not Optimal // Numer. Math. 13, p. 354-356, 1969.