УЧЕНЫЕ ЗАПИСКИ ПЕТРОЗАВОДСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА

Декабрь, № 8. Т. 2 **Физико-математические науки** **2012**

УДК 519.114

АНАСТАСИЯ МИХАЙЛОВНА ВАЛОВА
магистрант 2-го года обучения математического факультета, Петрозаводский государственный университет (Петрозаводск, Российская Федерация)
valova.a.m@gmail.com

SEARCH OF LINEAR RECURRENCE CORRELATION WITH CONSTANT INTEGER COEFFICIENTS IN PRE-SET SEQUENCE BY MEANS OF EUCLUDEAN ALGORIGHM*

Recurrences with properties mentioned in the title are common for physical problems which can be traced to the problems of enumerative combinatorics and then solved with the help of the transfer matrix method. We suggested a modification of Euclidean algorithm, which uses modular arithmetic to solve the problem.

Key words: Linear recurrence, Euclidean algorithm, Modular arithmetic

INTRODUCTION

Recurrences pointed in the title are usual for the enumerative combinatoric problems that can be solved by using the transfer matrix method [2], [3], [4], [5]. There are some software tools for finding such recurrences – the so called Computer Algebra Systems (CAS). However, the testing has shown that the ability of Mathematica 7.0/8.0 is not enough to get repeated well-known results [8]. CAS Maple 14 is more effective but the increase of the recurrence order dramatically raises the amount of RAM to be used. For example, the attempt of finding the recurrence with the order of 2086 for the number of Hamiltonian circuits in $P_{12} \times P_n$ [4] has failed the system.

The paper [1] was dedicated to the Dixon's algorithm application [7] to the problem of finding linear recurrence. It was extraordinary faster than the known software tools because of the p-adic expansions' usage. This paper shows that the Euclidean algorithm [6] modified in such a way that it uses modular arithmetic allows us to reduce the running time comparing to Dixon's algorithm in some cases.

ALGORITHM

According to the Cayley – Hamilton theorem any sequence $a_1, a_2, a_3 ...$ which has been obtained by use of transfer matrix method satisfies a linear homogeneous recurrence with the constant coefficients

$$a_k = \sum_{j=1}^{N} c_j \cdot a_{k-j}, k > N, \tag{1}$$

where N is known order. The coefficients of the recurrence c_j can be found out of the system of linear equations $\mathbf{A}_N \cdot \mathbf{c}_N = \mathbf{b}_N$ where

$$\mathbf{A}_N = \begin{pmatrix} a_N & a_{N-1} & \cdots & a_1 \\ a_{N+1} & a_N & \cdots & a_2 \\ \vdots & \vdots & \ddots & \vdots \\ a_{2N-1} & a_{2N-2} & \cdots & a_N \end{pmatrix}, \mathbf{c}_N = \begin{pmatrix} c_1 \\ c_1 \\ \vdots \\ c_N \end{pmatrix}, \mathbf{b}_N = \begin{pmatrix} a_{N+1} \\ a_{N+2} \\ \vdots \\ a_{2N} \end{pmatrix}.$$

Note that the uniqueness of the solution is ensured by the inequality in (1). The system requires first $2N$ terms of the sequence to be known. So our case is significantly different from the other known algorithms for recovering the recurrence for terms in the given sequences with unknown properties.

If N *is* unknown, then we have to consider its maximum value, assuming N to be equal to the order of the transfer matrix which can significantly exceed the real order of our relation. In the Dixon's algorithm we have to find the rank of $\mathbf{A}_N$ which shows the real order of the recurrence. The Euclidian algorithm works rather differently – it doesn't need to determine the order before calculations. Let's begin with its classical description [6].

Iterative scheme of the algorithm can be represented as follows. Let us assume

$$s^{(0)}(x) = x^{2N}, t^{(0)}(x) = \sum_{i=0}^{2N-1} a_i \cdot x^i, \text{ and}$$

$$A^{(0)}(x) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

then we repeat the following procedure until get $\deg t^{(r)}(x) < N - 1$

$$Q^{(r)}(x) = \left\lfloor \frac{s^{(r-1)}(x)}{t^{(r-1)}(x)} \right\rfloor$$

$$A^{(r)}(x) = \begin{bmatrix} 0 & 1 \\ 1 & -Q^{(r)}(x) \end{bmatrix} \cdot A^{(r-1)}(x)$$

$$\begin{bmatrix} s^{(r)}(x) \\ t^{(r)}(x) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & -Q^{(r)}(x) \end{bmatrix} \cdot \begin{bmatrix} s^{(r-1)}(x) \\ t^{(r-1)}(x) \end{bmatrix}.$$

After that, coefficients of the polynomial $f(x) = -\Delta^{(-1)} \cdot A_{22}^{(r)}(x)$ where $\Delta = A_{22}^{(r)}(0)$ correspond to the coefficients c_i of our recurrence (1).

However, the usage of rational numbers for coefficients of all involved polynomials brings extreme growth of numerator and denominator at the intermediate steps of the algorithm: the size of fractions have been grown up to one million of bits on hard tests. At the last step when calculating $f(x) = -\Delta^{(-1)} \cdot A_{22}^{(r)}(x)$, the coefficients should become integer and rather small. For example, the number of Hamiltonian cycles in $P_{10} \times P_n$ are subjected to the recurrence of order 346 [8] (sequence

A180504). The required initial numbers in the sequence of this numbers can exceed 1000 decimal digits, while the biggest coefficient c_j for the recurrence has only 70 decimal digits. Therefore, it was decided to use modular arithmetic to avoid hard fraction calculations.

Let us consider that all calculations are performed by modulo, where p is a prime integer. First, we consider some small prime number, for example $p = 2^{31} - 1$. Consider the modified Euclidean algorithm, where all polynomial operations are performed by modulo p. So, all coefficients of all polynomials at the intermediate steps of the algorithm are less than p. The modified algorithm will give us some solution $c_j^{(p)}$ by modulo p. Then we have to restore the right coefficients (if possible).

Consider that the numbers c_j lie in the range $(-\lfloor p/2 \rfloor, \lfloor p/2 \rfloor)$. So if a number $c_j^{(p)}$ is greater or equal to $\lfloor p/2 \rfloor$, then we have to decide $c_j = c_j^{(p)} - p$, otherwise $c_j = c_j^{(p)}$. On having done this, we should check the solution by substitution c_j into (1). If all the numbers a_k for $k = N, N+1, \ldots, 2N$ are satisfying the recurrence, we have found the right solution. Otherwise we have to greaten p. Good strategy is to find the smallest prime that is bigger than p^2 and to repeat all calculations again. Another strategy is to input the maximum number of bits b to the program. It finds the nearest to 2^b prime p and after having done all calculation shous if the answer is found or it isn't. On the one hand, using the second strategy the user have to select the upper bound on b by hands, while the first strategy will make all selections automatically. On the other hand, the second strategy allows to select more accurate bound on b, then the first one.

EXAMPLE

Let us consider an example: we have a sequence of integers, 1, 14, 154, 1696, 18684, 205832, 2267544, 24980352, 275195536, 3031685984, … This sequence corresponds to the number of Hamiltonian cycles in $P_5{\times}P_{2n}$ for $n = 1, 2, \ldots, 10$ [9] (sequence A006865). Let us apply the algorithm performing all operation by modulo $p=13$. Let us assume

$$N = 3,\ s^{(0)}(x) = x^6,$$

$$t^{(0)}(x) = 1 + x + 11x^2 + 6x^3 + 3x^4 + 5x^5, \text{ and}$$

$$A^{(0)}(x) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Begin calculating

$$Q(1)(x) = 9x + 4$$

$$A^{(1)}(x) = \begin{bmatrix} 0 & 1 \\ 1 & 4 + 9x \end{bmatrix}$$

$$s^{(1)}(x) = 1 + x + 11x^2 + 6x^3 + 3x^4 + 3x^5$$

$$t^{(1)}(x) = 9 + 12x^2 + 7x^3 + 12x^4$$

We will continue until the degree of the polynomial $t(x)$ is less than $N - 1 = 3 - 1 = 2$.

$$Q^{(2)}(x) = 10x + 2$$

$$A^{(2)}(x) = \begin{bmatrix} 1 & 4 + 9x \\ 11 + 3x & 9 + 6x + 12x^2 \end{bmatrix}$$

$$s^{(2)}(x) = 9 + 12x + 7x^3 + 12x^4$$

$$t^{(2)}(x) = 9 + 2x + 2x^3$$

$$Q^{(3)}(x) = 6x + 10$$

$$A^{(3)}(x) = \begin{bmatrix} 11 + 3x & 9 + 6x + 12x^2 \\ 8 + 8x + 8x^2 & 10 + 7x + 6x^3 \end{bmatrix}$$

$$s^{(3)}(x) = 9 + 2x + 2x^3$$

$$t^{(3)}(x) = 10 + 4x.$$

We see that the degree of polynomial $t^{(3)}(x)$ is equal to 1 and it is less than 2. So we have $f(x) = \frac{A_{22}(x)}{A_{22}(0)} \bmod p = 11x + 2x^3$ and respectively we have got the recurrence $a_n = 11a_{n-1} + 2a_{n-3}$. Having done this, we should check the solution by substitution c_j into (1). In our case all the numbers a_n for $n = deg\ Q + 1, \ldots, 2N$ are satisfying the recurrence, we can argue that the solution is right. If we chose $p=7$, for example, we would get $a_n = 4a_{n-1} + 2a_{n-3}$. Obviously, this is the wrong answer. That is why we have chosen another p.

TEST RESULTS

The testing was conducted on four tasks, the first, the second and the third problems are enumeration the number of Hamiltonian circles in rectangular lattices Pm × Pn [4], the number Hamiltonian circles in thin cylinder Cm × Pn [4] and the number Hamiltonian circles in torus Cm × Cn [4] and the fourth one is the placement of kings in a rectangular chessboard of size 2m × 2n [2].

Table 1

Lattice size, the corresponding order of the relation and the spent time (s) for the problem of finding the number of Hamiltonian circles in rectangular lattices

m	N	Mathematica 8	Maple 14	Dixon	Euclid
7	18	0,3	0,5	0,0	0,0
8	66	25,9	0,5	0,0	0,1
9	104	–	2,0	0,2	0,2
10	346	–	14,2	1,7	1,6
11	671	–	555,9	32,0	32,3
12	2086	–	–	821,8	650,0

Table 2

Cylinder size, the corresponding order of the relation and the spent time (s) for the problem of finding the number of Hamiltonian circles in thin cylinder

m	*N*	Mathematica 8	Maple 14	Dixon	Euclid
8	20	0,4	0,0	0,0	0,0
9	51	8,6	0,0	0,0	0,1
10	74	92,5	1,0	0,0	0,1
11	246	–	6,8	0,7	0,6
12	303	–	26,4	1,3	1,2
13	1320	–	–	197,1	160,5
14	1514	–	–	346,8	267,5

Table 3

Torus size, the corresponding order of the relation and the spent time(s) for the problem of finding the number of Hamiltonian circles in torus

m	*N*	Mathematica 8	Maple 14	Dixon	Euclid
4	28	0,4	0,0	0,0	0,088
5	84	74,4	0,6	0,0	0,113
6	257	–	26,3	0,7	0,492
7	856	–	430,8	29,5	19,612
8	2785	–	–	1849,4	>2 GiB RAM

We have used the computer with Intel Intel Core 2 Duo E-8400 @ 3.00 GHz, 4 GiB RAM, 32-bit operating system. Tables 1, 2, 3 and 4 show the runtime of Mathematica 8, Maple 14, Dixon's algorithm [1] and the described Euclidean algorithm. The runtime is given in seconds.

Table 4

Board size, the corresponding order of the relation and the spent time (s) for the problem of placement of kings in a rectangular chessboard

m	*N*	Mathematica 8	Maple 14	Dixon	Euclid
4	17	0,4	0,0	0,0	0,049
5	31	0,5	0,0	0,0	0,045
6	75	38,9	0,0	0,2	0,194
7	124	812,6	1,5	0,1	0,255
8	307	–	15,4	1,5	2,699
9	548	–	774,7	10,5	17,413
10	1318	–	–	455,1	>2 GiB RAM

As we can see that a described Euclidean algorithm is faster than other known tools in our test cases. Of course, one could test this algorithm in other test cases and get the other results. We had to show that in case of the leak of RAM Euclidean algorithm works better with exponentially growing integer sequences.

* Работа выполнена при поддержке Программы стратегического развития (ПСР) ПетрГУ в рамках реализации комплекса мероприятий по развитию научно-исследовательской деятельности на 2012–2016 гг.

REFERENCES

1. Валова А. М. Получение линейного рекуррентного соотношения с постоянными коэффициентами по заданной последовательности // Материалы XV Всероссийской науч.-практ. конференции «Научное творчество молодежи», 28–29 апреля 2011 г. Томск: Изд-во Томского ун-та, 2011. С. 48–51.
2. Караваев А. М. Задача о расстановке шахматных королей: материалы // Современные проблемы гуманитарных и естественных наук: Материалы II Междунар. науч.-практ. конф. М., 2010. Т. II. С. 12–16.
3. Караваев А. М. Усовершенствованный метод матрицы переноса для подсчета гамильтоновых цепей на прямоугольных решетках и цилиндрах // Информационные процессы 2011. Т. 11. № 3. С. 336–347.
4. Караваев А. М. Кодирование состояний в методе матрицы переноса для подсчета гамильтоновых циклов на прямоугольных решетках, цилиндрах и торах // Информационные процессы 2011. Т. 11. № 4. С. 476–499.
5. Караваев А. М. Подсчет предгамильтоновых циклов на семействах решеточных графов // Ученые записки Петрозаводского государственного университета Сер. «Естественные и технические науки». 2011. № 6(119). С. 97–102.
6. Blahut R. E. Fast Algorithms for Digital Signal Processing. Addison-Wesley, 1984. 448 p.
7. Dixon J. D. Exact Solution of Linear Equations Using p-adic Expansions // Numerische Mathematik 1982. Vol. 40. P. 137–141.
8. Stoyan R., Strehl V. Enumeration of Hamiltonian Circuits in Rectangular Grids // Journal of Combinatorial Mathematics and Combinatorial Computing 1996. Vol. 21. P. 109–127.
9. The On-Line Encyclopedia of Integer Sequences [Electronic resource]. Access mode: http://oeis.org