Strassen's Algorithm

Anindita Pal

Information Technology Department
Techno India College of Technology, WBUT UNIVERSITY
Newtown Rajarhat, Kolkata-700156

Abstract **- This abstract is based on the discussion about the easy technique of seven rules of Strassen's Matrix algorithm.**

I. INTRODUCTION

Strassen's matrix is based on the concepts of "Divide and Conquer Technique". Here I have focused on the tricks and techniques and sequence of the seven rules of Strassen's matrix Algorithm, taking two matrices A and B where product of the two matrices is C.

Let us consider A, B, C are (n/2)*(n/2) matrices. So that we can compute here C=AB

$$= \begin{pmatrix} c11 & c12 \\ c21 & c22 \end{pmatrix}$$

a11b11 + a12b21	**a11b12 + a12b22**
a21b11 + a22b21	**a21b12 + a22b22**

Before going to the discussion of the seven rules, we will see that there is a sequence lying between each pair like,[Q1],[Q2,Q5],[Q3,Q4],[Q6,Q7].So that it will be easy to remember those formulas.

II. DISCUSSION ABOUT THE SEVEN RULES

1. ***Let,Q1= (a11+a22) (b11+b22) =a11b11+a11b22+a22b11 +b22a22***
   actually this is basically a cross operation between two matrix......a11 is in the angular position of a22 and similarly b11 is in the angular position of b22......

   a11 a12 b11 b12
   ×
   a21 a22 b21 b22

2. ***Let, Q2 = (a21+a22) b11 = a21b11+a22b11***
3. ***Let, Q5 = (a11+a12) b22 = a11b22+a12b22***

   Here at first Q2 operation is happening with the second row of the first matrix multiplied by first element of first row, first column of second matrix. Similarly, reverse will be happening in case of Q5. i. e the first row of the first matrix will be multiplied by 4th element of second row, second column of second matrix. As both of these operations are row wise so arithmetical operation should be "+ "......

   a11 a12 b11 b12
   ×
   a21 a22 b21 b22

4. ***Let, Q3= (a11) (b12-b22)*** **= a11b12-*a11b22***
5. ***Let, Q4= (a22) (b21-b11) = a22b21-a22b11***

   Here at first Q3 operation is happening with the second column of the second matrix multiplied by first element of first row, first column of first matrix. Similarly, reverse will be happening in case of Q4. i.e. The first column of the second matrix will be multiplied by 4th element of second row, second column of first matrix. As both of these operations are column wise so arithmetical operation should be "minus "......

   a11 a12 b11 b12
   ×
   a21 a22 b21 b22

6. ***Let,Q6 = (a21-a11) (b11+b12) = a21b11+a21b12-a11b11 - a11b12***
7. ***Let,Q7= (a12-a22) (b21+b22) = a12b21+a12b22-a22b21 - a22b22***

Here at first Q6 operation is happening with the first column of first matrix multiplied by first row of the second matrix. Similarly, reverse will be happening in case of Q7. i.e. the second column of the first matrix will be multiplied by second row of the second matrix. Similarly here row and column operation of both the matrix is occurring. so arithmetical operation “add” and “minus” will be happening.

a11 a12 × b11 b12
a21 a22 b21 b22

So we can conclude from the above equations that

For

1. [Q2, Q5] the product will be in the form of addition,as we are performing row operation mostly.
2. For [Q3, Q4] the product will be in form of subtractions as we are performing column wise operations mostly.
3. For [Q6, Q7], row-column both operations are being performed, so the products will be in form of additions and subtractions.

Now the products will be:

C11

= Q1 + Q4 - Q5 +Q7
= a11b11+a11b22+a22b11+b22a22+b21a22-a22b11+a12b21+a12b22-a22b21-a22b22-a11b22-a12b22
= (a11b11 + a12b21)

C22

= Q1 +Q3 - Q2 + Q6
= a11b11+a11b22+a22b11+b22a22 +a11b12-a11b22-a21b11-a22b11 +a21b11+a21b12-a11b11-a11b12
= (a21b12 + a22b22)

C12

= Q3 + Q5
= a11b12-a11b22+a11b22+a12b22
= (a11b12 + a12b22)

C21

= Q4 + Q2
= a22b21-a22b11+a21b11+a22b11
= (a21b11 + a22b21)

So it satisfy the condition i.e

$$= \begin{pmatrix} c11 & c12 \\ c21 & c22 \end{pmatrix}$$

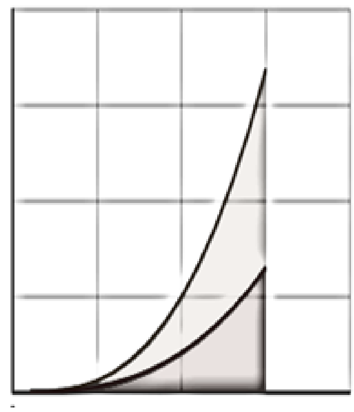
a11b11 + a12b21	**a11b12 + a12b22**
a21b11 + a22b21	**a21b12 + a22b22**

So our Strassen’s algorithm will be like this:

Strassen (A, B)

1. If n = 1 Output A × B
2. Else
3. Compute a11, b11. . . a22, b22 % by computing m = n/2
4. Q1 ← Strassen(a11+a22,b11+b22)
5. Q2 ← Strassen(a21+a22,b11)
6. Q3 ← Strassen(a11,b12-b22)
7. Q4 ← Strassen(a22,b21-b11)
8. Q5 ← Strassen(a11+a12,b22)
9. Q6 ← Strassen(a21-a11,b11+b12)
10. Q7 ← Strassen(a12-a22,b21+b22)
11. C11← Q1+Q4-Q5+Q7
12. C22← Q1+Q3-Q2+Q6
13. C12 ← Q3+Q5

14. C21 ←Q4+Q2
15. Output C
16. End If



Time complexity of this algorithm will be

$T(n) = \{ m \quad if \; n = 1$

7 T (n/2) +18(n/2) ^2a if n>=2

={ $m \quad if \; n = 1$

7 T (n/2) + (9a/2) n^2 if n>=2

Assuming that n is a power of 2

T (n)

=[m+ [{(9a/2)2^2}/(7-2^2)]] n^log7- [{(9a/2)2^2}/(7-2^2)]n^2

=m (n^log7)+6a(n^log7)-6a(n^2)

= Θ (n^log7)

III. COMPARATIVE ANALYSIS OF O(N^3)

The definition of matrix multiplication is only valid if the width of first matrix equals the height of the second matrix. Let us consider for matrix A with dimensions m by n, and B with dimensions n by p, the result of A multiply B is an m by p matrix, where the elements of AB are given by:

$$(AB)_{i,j} = \sum_{r=1}^{n} A_{i,r} B_{r,j}$$

A block matrix is a partition of a matrix into smaller matrices, and each block itself is treated as a new element of the original matrix. For example, the matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}$$

can be split into four 2 by 2 block matrices

$$A_{11} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad A_{12} = \begin{bmatrix} a_{13} & a_{14} \\ a_{23} & a_{24} \end{bmatrix} \quad A_{21} = \begin{bmatrix} a_{31} & a_{32} \\ a_{41} & a_{42} \end{bmatrix} \quad A_{22} = \begin{bmatrix} a_{33} & a_{34} \\ a_{43} & a_{44} \end{bmatrix}$$

Then matrix

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

The product of block matrices can be computed using the principle of conventional matrix multiplication. The computation follows the same rule, which means the width of first partitioned matrix equals the height of the second partitioned matrix. For example, for an m by n matrix A with partitioned dimensions a by b, and a n by p matrix B with partitioned dimensions b by c, the product of two partitioned matrices C can be calculated by:

$$C_{i,j} = \sum_{r=1}^{n} A_{i,r} B_{r,j}$$

$$\begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \cdot \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

The standard matrix multiplication takes approximately 2N3 (where N = 2n) arithmetic operations (additions and multiplications); the asymptotic complexity is O(N3).

The number of additions and multiplications required in the Strassen algorithm can be calculated as follows: let f(n) be the number of operations for a 2n × 2n matrix. Then by recursive application of the Strassen algorithm, we see that f(n) = 7f(n-1) + l4n, for some constant l that depends on the number of additions performed at each application of the algorithm. Hence f(n) = (7 + o(1))n, i.e., the asymptotic complexity for multiplying matrices of size N = 2n using the Strassen algorithm is

$$O([7+o(1)]^n) = O(N^{\log_2 7+o(1)}) \approx O(N^{2.8074}).$$

The reduction in the number of arithmetic operations however comes at the price of a somewhat reduced numerical stability and the algorithm also requires significantly more memory compared to the naive algorithm. Both initial matrices must have their dimensions expanded to the next power of 2, which results in storing up to four times as many elements, and the seven auxiliary matrices each contain a quarter of the elements in the expanded ones.

IV. Conclusion

However it can be concluded that by applying MASTERS THEOREM, the time complexity of this row-column multiplication will be nearly equal to O(N^2.808) or O(N^3).

Application

It can be said that Strassen matrix algorithm can be used to solve the problem involving large matrices.

V. Acknowledgement

I am grateful to my mentor Mr. Ayan Chakraborty (Assistant professor of Techno India College of Technology, IT dept.) and the whole IT faculty for inspiring me a lot and helping me In clearing my concepts.

References

[1] http://mathworld.wolfram.com/StrassenFormulas.html
[2] http://en.wikipedia.org/wiki/Strassen_algorithm
[3] https://software.intel.com/sites/default/files/m/c/d/5/3/d/24469-Strassen_akki.pdf
[4] http://www.geeksforgeeks.org/strassens-matrix-multiplication/
[5] http://www.cs.mcgill.ca/~pnguyen/251F09/matrix-mult.pdf
[6] http://www.stoimen.com/blog/2012/11/26/computer-algorithms-strassens-matrix-multiplication/
[7] http://www.leda-tutorial.org/en/official/ch02s02s03.html
[8] http://www.ncs.org.ng/wp-content/uploads/2011/08/ITePED2011-Paper15.pdf
[9] http://en.wikipedia.org/wiki/Time_complexity#Polynomial_time
[10] D. Bini, M. Capovani, F. Romani, and G. Lotti. O(n2.7799) complexity for n_n approximate matrix multiplication. Inf. Process. Lett., 8(5):234–235, 1979.
[11] Juby Mathew, Dr. R Vijaya Kumar, Comparative Study of Strassen's Matrix Multiplication, IJCST Vol. 3, Issue 1, Jan. - March 2012
[12] Sara Robinson,"Toward an Optimal Algorithm for Matrix Multiplication", From SIAM News, Vol. 38, No. 9, November 2005. [Online] Available: http://www.siam.org/pdf/news/174.pdf>