Strassen's Algorithm for Matrix Multiplication

Joe Song

February 27, 2020

The lecture notes are mostly based on Section 4.2 of Cormen, Leiserson, Rivest, and Stein. Introduction to Algorithms. 3rd Ed. 2009. MIT Press. Cambridge, Massachusetts.

Contents

1 Running time of matrix multiplication

Goal: To multiply two $n \times n$ matrices A and B

$$C = AB$$

using $o(n^3)$ time (not big-O).

$$C = \begin{pmatrix} r & s \\ t & u \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} e & f \\ g & h \end{pmatrix} = A\,B$$

$$\begin{aligned} r &= ae + bg \\ s &= af + bh \\ t &= ce + dg \\ u &= cf + dh \end{aligned}$$

Implementing the above by divide-and-conquer, we have

$$T(n) = 8T(n/2) + \Theta(n^2)$$

which is $\Theta(n^3)$, not $o(n^3)$.

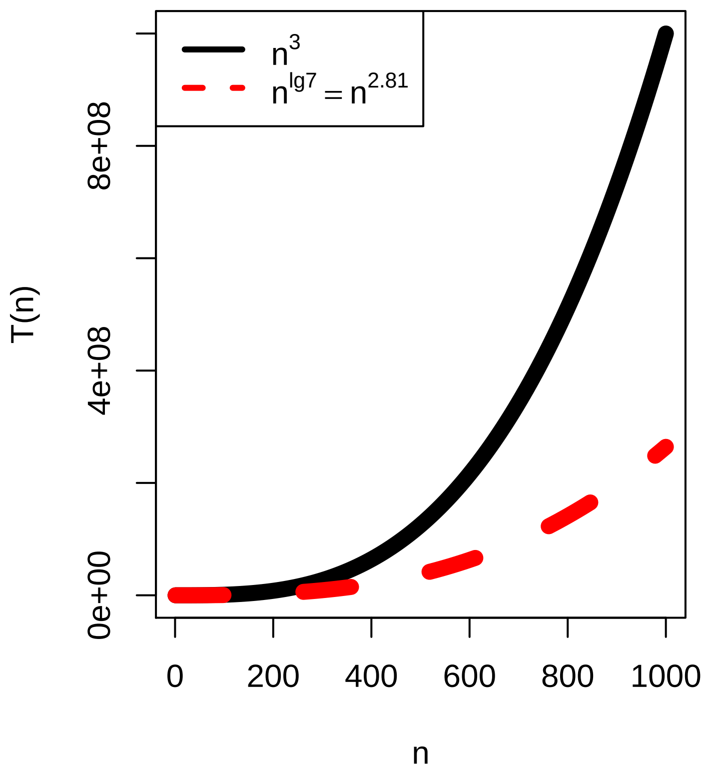
2 Description of Strassen's algorithm

Strassen method uses only

- 7 recursive multiplications of two $n/2 \times n/2$ matrices, and
- $\Theta(n^2)$ scalar additions and subtractions

yielding the recurrence

$$T(n) = 7T(n/2) + \Theta(n^2) = \Theta(n^{\lg 7}) \approx \Theta(n^{2.81})$$



Four steps:

1. Divide the input $n \times n$ matrix to four $n/2 \times n/2$ sub-matrices
2. Compute 14 $n/2 \times n/2$ intermediate sub-matrices $A_1, \ldots, A_7$ and $B_1, \ldots, B_7$ by 10 sub-matrix additions or subtractions.
3. Compute **seven** $n/2 \times n/2$ matrix products $P_1 = A_1 B_1, \ldots, P_7 = A_7 B_7$
4. Compute sub-matrices r, s, t, u in C by 8 sub-matrix additions or subtractions among P_i.

$$P_i = A_i B_i = (\alpha_{i1} a + \alpha_{i2} b + \alpha_{i3} c + \alpha_{i4} d) \cdot (\beta_{i1} e + \beta_{i2} f + \beta_{i3} g + \beta_{i4} h)$$

$$\begin{aligned} r &= ae + bg \\ &= (a\ b\ c\ d) \begin{pmatrix} +1 & 0 & 0 & 0 \\ 0 & 0 & +1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} e \\ f \\ g \\ h \end{pmatrix} \\ &= \begin{array}{ccccc} & e & f & g & h \\ a & + & \cdot & \cdot & \cdot \\ b & \cdot & \cdot & + & \cdot \\ c & \cdot & \cdot & \cdot & \cdot \\ d & \cdot & \cdot & \cdot & \cdot \end{array} \end{aligned}$$

"+": for +1

"−": for −1

"·": for 0

$$s = af + bh = \begin{array}{ccccc} & e & f & g & h \\ a & \cdot & + & \cdot & \cdot \\ b & \cdot & \cdot & \cdot & + \\ c & \cdot & \cdot & \cdot & \cdot \\ d & \cdot & \cdot & \cdot & \cdot \end{array}$$

$$t = ce + dg = \begin{array}{ccccc} & e & f & g & h \\ a & \cdot & \cdot & \cdot & \cdot \\ b & \cdot & \cdot & \cdot & \cdot \\ c & + & \cdot & \cdot & \cdot \\ d & \cdot & \cdot & + & \cdot \end{array}$$

$$u = cf + dh = \begin{array}{ccccc} & e & f & g & h \\ a & \cdot & \cdot & \cdot & \cdot \\ b & \cdot & \cdot & \cdot & \cdot \\ c & \cdot & + & \cdot & \cdot \\ d & \cdot & \cdot & \cdot & + \end{array}$$

$$\begin{aligned} P_1 &= A_1 B_1 \\ &= a \cdot (f - h) \\ &= af - ah \\ &= \begin{array}{ccccc} & e & f & g & h \\ a & \cdot & + & \cdot & - \\ b & \cdot & \cdot & \cdot & \cdot \\ c & \cdot & \cdot & \cdot & \cdot \\ d & \cdot & \cdot & \cdot & \cdot \end{array} \end{aligned}$$

$$
\begin{aligned}
P_2 &= A_2 B_2 \\
&= (a+b)\cdot h \\
&= ah + bh \\
&= \begin{array}{ccccc} & e & f & g & h \\ a & \cdot & \cdot & \cdot & + \\ b & \cdot & \cdot & \cdot & + \\ c & \cdot & \cdot & \cdot & \cdot \\ d & \cdot & \cdot & \cdot & \cdot \end{array}
\end{aligned}
$$

$\implies s = P_1 + P_2$

$$
\begin{aligned}
P_3 &= A_3 B_3 \\
&= (c+d)\cdot e \\
&= ce + de \\
&= \begin{array}{ccccc} & e & f & g & h \\ a & \cdot & \cdot & \cdot & \cdot \\ b & \cdot & \cdot & \cdot & \cdot \\ c & + & \cdot & \cdot & \cdot \\ d & + & \cdot & \cdot & \cdot \end{array}
\end{aligned}
$$

$$
\begin{aligned}
P_4 &= A_4 B_4 \\
&= d \cdot (g - e) \\
&= dg - de \\
&= \begin{array}{ccccc}
 & e & f & g & h \\
a & \cdot & \cdot & \cdot & \cdot \\
b & \cdot & \cdot & \cdot & \cdot \\
c & \cdot & \cdot & \cdot & \cdot \\
d & - & \cdot & + & \cdot
\end{array}
\end{aligned}
$$

$\Longrightarrow\ t = P_3 + P_4$

$$
\begin{aligned}
P_5 &= A_5 B_5 \\
&= (a + d) \cdot (e + h) \\
&= ae + ah + de + dh \\
&= \begin{array}{ccccc}
 & e & f & g & h \\
a & + & \cdot & \cdot & + \\
b & \cdot & \cdot & \cdot & \cdot \\
c & \cdot & \cdot & \cdot & \cdot \\
d & + & \cdot & \cdot & +
\end{array}
\end{aligned}
$$

$$
\begin{aligned}
P_5 + P_4 - P_2 &= ae + dh + dg - bh \\
&= \begin{array}{ccccc}
 & e & f & g & h \\
a & + & \cdot & \cdot & + \\
b & \cdot & \cdot & \cdot & - \\
c & \cdot & \cdot & \cdot & \cdot \\
d & \cdot & \cdot & + & +
\end{array}
\end{aligned}
$$

$$
\begin{aligned}
P_6 &= A_6 B_6 \\
&= (b-d)\cdot(g+h) \\
&= bg + bh - dg - dh \\
&= \begin{array}{c|cccc}
 & e & f & g & h \\
a & \cdot & \cdot & \cdot & \cdot \\
b & \cdot & \cdot & + & + \\
c & \cdot & \cdot & \cdot & \cdot \\
d & \cdot & \cdot & - & -
\end{array}
\end{aligned}
$$

$$
\implies r = P_5 + P_4 - P_2 + P_6 = ae + bg = \begin{array}{c|cccc}
 & e & f & g & h \\
a & + & \cdot & \cdot & \cdot \\
b & \cdot & \cdot & + & \cdot \\
c & \cdot & \cdot & \cdot & \cdot \\
d & \cdot & \cdot & \cdot & \cdot
\end{array}
$$

$$
\begin{aligned}
P_5 + P_1 - P_3 &= ae + af - ce + dh \\
&= \begin{array}{c|cccc}
 & e & f & g & h \\
a & + & + & \cdot & \cdot \\
b & \cdot & \cdot & \cdot & - \\
c & - & \cdot & \cdot & \cdot \\
d & \cdot & \cdot & \cdot & +
\end{array}
\end{aligned}
$$

$$
\begin{aligned}
P_7 &= A_7 B_7 \\
&= (a-c)\cdot(e+f) \\
&= ae + af - ce - cf \\
&= \begin{array}{ccccc} & e & f & g & h \\ a & + & + & \cdot & \cdot \\ b & \cdot & \cdot & \cdot & \cdot \\ c & - & - & \cdot & \cdot \\ d & \cdot & \cdot & \cdot & \cdot \end{array}
\end{aligned}
$$

$$
\Longrightarrow u = P_5 + P_1 - P_3 - P_7 = cf + dh = \begin{array}{ccccc} & e & f & g & h \\ a & \cdot & \cdot & \cdot & \cdot \\ b & \cdot & \cdot & \cdot & \cdot \\ c & \cdot & + & \cdot & \cdot \\ d & \cdot & \cdot & \cdot & + \end{array}
$$

Thus we use seven matrix products to compute $P_1, P_2, \ldots, P_7$. All remaining operations are additions.

3 Summary

$$
\begin{aligned}
P_1 &= a \cdot (f - h) \\
P_2 &= (a + b) \cdot h \\
P_3 &= (c + d) \cdot e \\
P_4 &= d \cdot (g - e) \\
P_5 &= (a + d) \cdot (e + h) \\
P_6 &= (b - d) \cdot (g + h) \\
P_7 &= (a - c) \cdot (e + f) \\
r &= P_5 + P_4 - P_2 + P_6 \\
s &= P_1 + P_2 \\
t &= P_3 + P_4 \\
u &= P_5 + P_1 - P_3 - P_7
\end{aligned}
$$

There are:

- Seven (7) multiplications of two $n/2 \times n/2$ submatrices
- Eighteen (18) additions or subtractions of two $n/2 \times n/2$ submatrices, leading to a total number of

  $$18 \times (n/2) \times (n/2) = \frac{9n^2}{2}$$

  operations

Thus

$$T(n) = 7T(n/2) + \frac{9n^2}{2}$$

giving rise to

$$T(n) = \Theta(n^{\lg 7})$$