Note for Strassen Algorithm

Xiaoyan He

2023-02-22

1 Ideas of Basic Matrix Multiplication

The idea behind Strassen algorithm is in the formulation of matrix multiplication as a recursive problem. Assume that we have two matrices in $\mathbb{R}^{n\times n}$ A and B. We can write A and B as follows:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \tag{1}$$

$$B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \tag{2}$$

Then we can write the matrix multiplication as follows:

$$AB = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{bmatrix} \tag{3}$$

However, we still need to do 8 matrix multiplications to get the result. The time complexity is $O(n^3)$.

2 Strassen Algorithm

The idea of Strassen algorithm is to reduce the number of matrix multiplications. We can write the matrix multiplication as follows:

$$M_1 = (A_{11} + A_{22})(B_{11} + B_{22}) \tag{4}$$

$$M_2 = (A_{21} + A_{22})B_{11} \tag{5}$$

$$M_3 = A_{11}(B_{12} - B_{22}) \tag{6}$$

$$M_4 = A_{22}(B_{21} - B_{11}) \tag{7}$$

$$M_5 = (A_{11} + A_{12})B_{22} \tag{8}$$

$$M_6 = (A_{21} - A_{11})(B_{11} + B_{12}) \tag{9}$$

$$M_7 = (A_{12} - A_{22})(B_{21} + B_{22}) \tag{10}$$

Then we can write the matrix multiplication $AB = C$ as follows:

$$AB = C = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} \tag{11}$$

$$C_{11} = M_1 + M_4 - M_5 + M_7 \tag{12}$$

$$C_{12} = M_3 + M_5 \tag{13}$$

$$C_{21} = M_2 + M_4 \tag{14}$$

$$C_{22} = M_1 - M_2 + M_3 + M_6 \tag{15}$$

We can see that we only need to do 7 matrix multiplications to get the result. The time complexity is $O(n^{\log_2 7})$. In general, we will recursively do the matrix multiplication until the size of the matrix is 1×1. Then we can use the basic matrix multiplication to get the result.

3 Limitation of Strassen Algorithm

Strassen's algo is not quite as numerically stable as the naive method. Strassen's algorithm only concerns dense matrices and that the matrices need to be fairly large for it to be worthwhile.

4 Coppersmith and Winograd's Algorithms