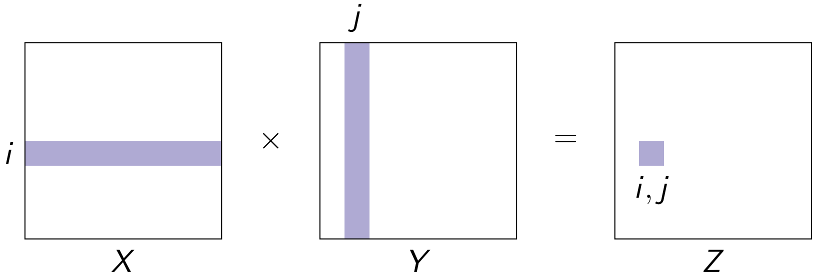
«Разделяй и властвуй»: умножение матриц

Александр Куликов

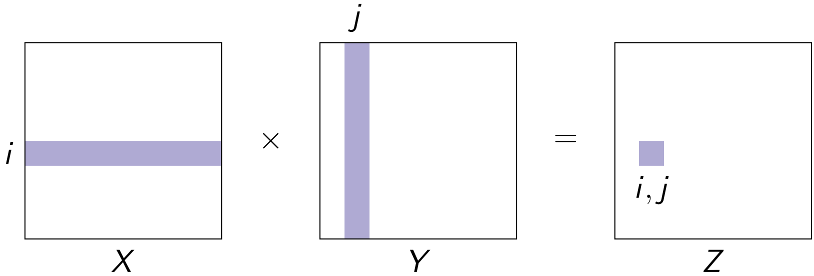
Онлайн-курс «Алгоритмы: теория и практика. Методы»
http://stepic.org/217

Умножение матриц



$$Z[i, j] = \sum_{k=1}^{n} X[i, k] \cdot Y[k, j]$$

Умножение матриц



$$Z[i,j] = \sum_{k=1}^{n} X[i,k] \cdot Y[k,j]$$

Время работы наивного алгоритма: $O(n^3)$.

Умножение матриц размера 2×2

$$X = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} \quad Y = \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix}$$

$$XY = \begin{bmatrix} X_{11}Y_{11} + X_{12}Y_{21} & X_{11}Y_{21} + X_{12}Y_{22} \\ X_{21}Y_{11} + X_{22}Y_{21} & X_{21}Y_{21} + X_{22}Y_{22} \end{bmatrix}$$

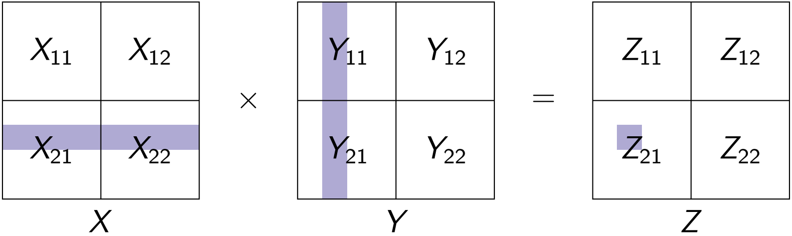
Умножение матриц размера 2×2

$$X = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} \quad Y = \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix}$$

$$XY = \begin{bmatrix} X_{11}Y_{11} + X_{12}Y_{21} & X_{11}Y_{21} + X_{12}Y_{22} \\ X_{21}Y_{11} + X_{22}Y_{21} & X_{21}Y_{21} + X_{22}Y_{22} \end{bmatrix}$$

Данная формула верна и в случае, когда X и Y — это матрицы размера $n \times n$, а X_{ij}, Y_{ij} — это их подматрицы размера $n/2 \times n/2$!

Визуальное пояснение



$$Z_{21} = X_{21}Y_{11} + X_{22}Y_{21}$$

Рекурсивный алгоритм со временем работы $O(n^3)$

$$XY = \begin{bmatrix} X_{11}Y_{11} + X_{12}Y_{21} & X_{11}Y_{21} + X_{12}Y_{22} \\ X_{21}Y_{11} + X_{22}Y_{21} & X_{21}Y_{21} + X_{22}Y_{22} \end{bmatrix}$$

$$T(n) = 8T(n/2) + O(n^2)$$

$$T(n) = O(n^3)$$

Алгоритм Штрассена

$$XY = \begin{bmatrix} X_{11}Y_{11} + X_{12}Y_{21} & X_{11}Y_{12} + X_{12}Y_{22} \\ X_{21}Y_{11} + X_{22}Y_{21} & X_{21}Y_{12} + X_{22}Y_{22} \end{bmatrix}$$

$$XY = \begin{bmatrix} P_5 + P_4 - P_2 + P_6 & P_1 + P_2 \\ P_3 + P_4 & P_1 + P_5 - P_3 - P_7 \end{bmatrix}$$

$$\begin{aligned} P_1 &= X_{11}(Y_{21} - Y_{22}) & P_5 &= (X_{11} + X_{22})(Y_{11} + Y_{22}) \\ P_2 &= (X_{11} + X_{12})Y_{22} & P_6 &= (X_{12} - X_{22})(Y_{21} + Y_{22}) \\ P_3 &= (X_{21} + X_{22})Y_{11} & P_7 &= (X_{11} - X_{21})(Y_{11} + Y_{12}) \\ P_4 &= X_{22}(Y_{21} - Y_{11}) & & \end{aligned}$$

Лемма

Время работы алгоритма Штрассена есть $O(n^{\log_2 7}) = O(n^{2.807\ldots})$.

Доказательство

- Рекуррентное соотношение: $T(n) = 7T(n/2) + O(n^2)$.
- В дереве рекурсии будет $\log_2 n$ уровней. На i-м уровне будет 7^i задач размера $n/2^i$, на каждую из которых уйдёт время $c(n/2^i)^2$.
-

$$\sum_{i=0}^{\log_2 n} 7^i c\left(\frac{n}{2^i}\right)^2 = cn^2 \sum_{i=0}^{\log_2 n} \frac{7^i}{4^i} = n^2 \cdot \Theta\left(\frac{7^{\log_2 n}}{4^{\log_2 n}}\right) = \Theta(n^{\log_2 7})$$

$\square$