

Algoritmo de Strassen

Diseño y análisis de algoritmos

Funcionamiento

- Funciona para matrices de $n \times n$ con $n=2^k$.
- En caso de que se quiera multiplicar una matriz de diferente tamaño se rellenan las filas y/o columnas con ceros para llegar a una de $n \times n$ y $n=2^k$.

k	n	$n \times n$
0	1	1x1
1	2	2x2
2	4	4x4
3	8	8x8
4	16	16x16
5	32	32x32
6	64	64x64
7	128	128x128
8	256	256x256
9	512	512x512
10	1024	1024x1024

Funcionamiento

- Suponiendo que tenemos dos matrices A y B de nxn con n=2, entonces una forma de obtener la matriz C resultado de la multiplicación de A y B es:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \times \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}$$

$$m_1 = (a_{11}+a_{22})(b_{11}+b_{22})$$

$$m_2 = (a_{21}+a_{22})b_{11}$$

$$m_3 = a_{11}(b_{12}-b_{22})$$

$$m_4 = a_{22}(b_{21}-b_{11})$$

$$m_5 = (a_{11}+a_{12})b_{22}$$

$$m_6 = (a_{21}-a_{11})(b_{11}+b_{12})$$

$$m_7 = (a_{12}-a_{22})(b_{11}+b_{22})$$

$$C = \begin{pmatrix} m_1 + m_4 - m_5 + m_7 & m_3 + m_5 \\ m_2 + m_4 & m_1 + m_3 - m_2 + m_6 \end{pmatrix}$$

Este método hace solo
7 multiplicaciones y 18 sumas

Funcionamiento

- Ahora suponiendo dos matrices A y B de nxn con n=4

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

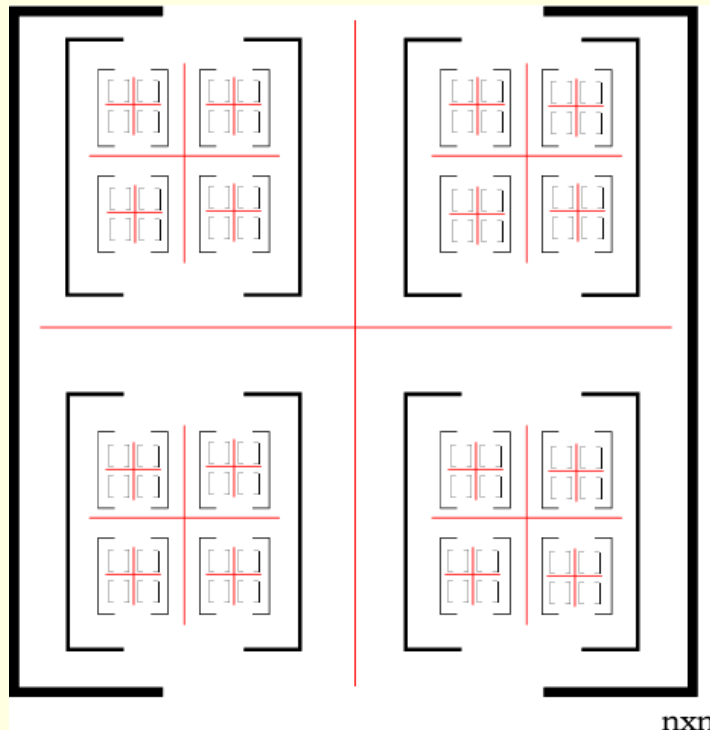
Entonces una forma de obtener la matriz C resultado de la multiplicación de A y B es:

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

The diagram illustrates the Strassen algorithm's approach to matrix multiplication. It shows a 4x4 matrix A partitioned into four 2x2 submatrices: A11 (top-left), A12 (top-right), A21 (bottom-left), and A22 (bottom-right). Similarly, matrix B is partitioned into B11, B12, B21, and B22. The resulting matrix C is also partitioned into C11, C12, C21, and C22. Red text labels these submatrices in the diagram.

Funcionamiento

- Dividir cada una de las matrices en 4 submatrices, y resolver por el método de Strassen el problema.



Cada submatriz
es de $n/2 \times n/2$

Funcionamiento

$$M1 = (A11+A22)(B11+B22)$$

$$M2 = (A21+A22)B11$$

$$M3 = A11(B12-B22)$$

$$M4 = A22(B21-B11)$$

$$M5 = (A11+A12)B22$$

$$M6 = (A21-A11)(B11+B12)$$

$$M7 = (A12-A22)(B11+B22)$$

$$M_1 = \left[\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} + \begin{pmatrix} a_{33} & a_{34} \\ a_{43} & a_{44} \end{pmatrix} \right] X \left[\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} + \begin{pmatrix} b_{33} & b_{34} \\ b_{43} & b_{44} \end{pmatrix} \right]$$

$$M_1 = \begin{pmatrix} a_{11} + a_{33} & a_{12} + a_{34} \\ a_{21} + a_{43} & a_{22} + a_{44} \end{pmatrix} X \begin{pmatrix} b_{11} + b_{33} & b_{12} + b_{34} \\ b_{21} + b_{43} & b_{22} + b_{44} \end{pmatrix}$$

$$C = \begin{pmatrix} C_{11} = M_1 + M_4 - M_5 + M_7 & C_{12} = M_3 + M_5 \\ C_{21} = M_2 + M_4 & C_{22} = M_1 + M_3 - M_2 + M_6 \end{pmatrix}$$

Ventajas del algoritmo

- La solución tradicional en matrices 2×2 , requiere de 8 multiplicaciones y 4 sumas...
- La solución de Strassen requiere de 7 multiplicaciones y 18 sumas/restas...
- Aparentemente, no es significativo el beneficio, pero ahorrar una multiplicación de matrices en el algoritmo de Strassen, a costa de más sumas o restas de matrices, tiene repercusiones significativas...

Costo del algoritmo

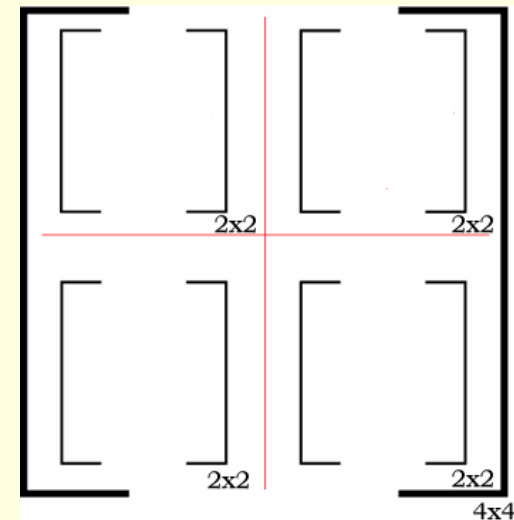
- Sea $T(n)$ el número total de operaciones (sumas, restas, multiplicaciones) que necesita el algoritmo para multiplicar dos matrices A y B de tamaño $n \times n$.

Tenemos que:

$$T(1)=1$$

$$T(2)=7+18=25$$

$$T(4)=7*25+18*4$$



7 multiplicaciones por cada submatriz de 2x2 con $T(2)$

18 sumas por cada submatriz

Costo del algoritmo

Desarrollando:

$$T(1)=1$$

$$T(2)=7+18=25$$

$$T(4)=7*25+18*4=247$$

$$T(8)=7*247+18*16=2017$$

$$T(16)=7*2017+18*64=15271$$

.

.

.

$$T(n)=7*T(n/2)+18*(n/2)^2$$

Número de submatrices



Tenemos que para calcular $T(n)$ se necesita recursividad. Sabemos que $n=2^k$ por lo que para $T(n)=T(2^k)$ se necesita $T(2^{(k-1)})=T(2^{(k)}/2)=T(n/2)$

Para las sumas tenemos $18*X$, X se Obtiene así:

$$T(2)=\dots+18*4^0$$

$$T(4)=\dots+18*4^1$$

$$T(8)=\dots+18*4^2$$

.

$$T(n)=\dots+18*4^{(k-1)}$$

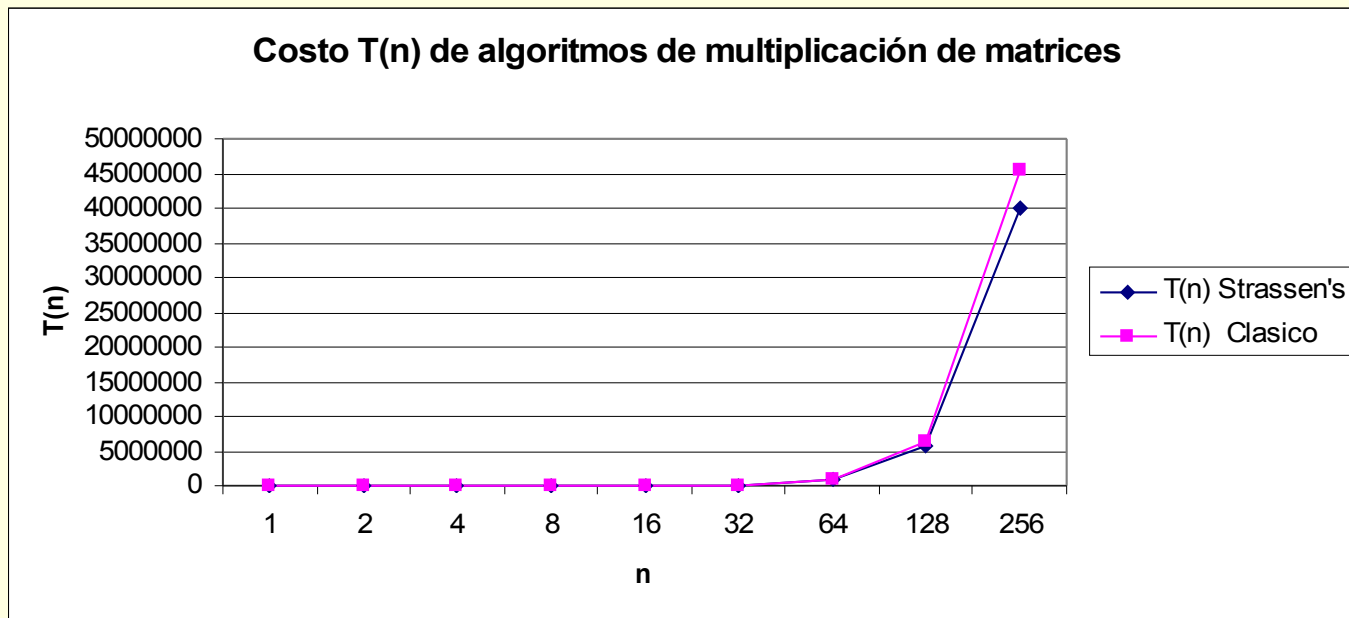
$$X=4^{(k-1)}=(2^2)^{(k-1)}=2^{(2*k-2)}=(2^k)*(2^k)/2^2=(n*n)/2^2=(n/2)^2$$

Costo del algoritmo

Para conocer que si el algoritmo de Strassen es menos costoso que el Algoritmo clásico, se obtuvo $T(n)$ de este último:

$$T(n) = 8T(n/2) + 4 \cdot (n/2)^2$$

En el siguiente grafico se pueden ver las comparaciones.



Costo del algoritmo

Tabla de resultados:

k	n	nxn	T(n) Strassen's	T(n) Clasico	Mejor algoritmo
0	1	1x1	1	1	Igual
1	2	2x2	25	12	Clasico
2	4	4x4	247	216	Clasico
3	8	8x8	2017	2040	Strassen
4	16	16x16	15271	16392	Strassen
5	32	32x32	111505	123192	Strassen
6	64	64x64	798967	896136	Strassen
7	128	128x128	5666497	6408120	Strassen
8	256	256x256	39960391	45397512	Strassen
9	512	512x512	280902385	319945272	Strassen
10	1024	1024x1024	1971035287	2248267656	Strassen

Aplicación del Teorema Maestro

Calcular C, a partir de P1, . . . , P7, en (n^2) operaciones.

Sea

$$A1 = a \quad B1 = g - h \quad P1 = ag - ah \quad T(n/2) + \Theta(n^2)$$

$$A2 = a + b \quad B2 = h \quad P2 = ah + bh \quad T(n/2) + \Theta(n^2)$$

$$A3 = c + d \quad B3 = e \quad P3 = ce + de \quad T(n/2) + \Theta(n^2)$$

$$A4 = d \quad B4 = f - e \quad P4 = df - de \quad T(n/2) + \Theta(n^2)$$

$$A5 = a + d \quad B5 = e + h \quad P5 = ae + ah + de + dh \quad T(n/2) + \Theta(n^2)$$

$$A6 = b - d \quad B6 = f + h \quad P6 = bf + bh - df - dh \quad T(n/2) + \Theta(n^2)$$

$$A7 = a - c \quad B7 = e + g \quad P7 = ae + ag - ce - cg \quad T(n/2) + \Theta(n^2)$$

$$7T(n/2) + \Theta(n^2)$$

Aplicación del Teorema Maestro

$$T(n) = 7T(n/2) + \Theta(n^2)$$

$$a=7; b=2; f(n)= \Theta(n^2) \text{ y } \text{nlg}_{ab}$$

Sustituyendo: $\text{nlg}_2 7$

Por lo tanto $f(n) < \text{nlg}_2 7$

Por lo que: $f(n) = O(n^{2.80})$

Aplicación del Teorema Maestro

Observaciones:

- Este algoritmo en la práctica solo es útil si las matrices son grandes y densas.
- Usando técnicas mas avanzadas se han encontrado mejores algoritmos. De hecho hay uno de orden $\Theta(n^{2.376})$.
- El algoritmo de Strassen ha tenido una repercusión fundamentalmente teórica, ya que es un algoritmo difícil de implementar.
- Es menos estable que el algoritmo clásico, ya que para errores similares en los datos produce mayores errores en el resultado.