

# Quanto è difficile moltiplicare matrici?

Complessità bilineare e rango di bordo

Federico Poloni

13 Aprile 2006

*"I see the eigenvalue in thine eye,  
I hear the tender tensor in thy sigh."  
—Stanislaw Lem, Cyberiad*

## L'algoritmo di Strassen

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

$$m_1 = (a_{12} - a_{22})(b_{21} + b_{22})$$

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

$$m_3 = (a_{11} - a_{21})(b_{11} + b_{12})$$

$$m_4 = (a_{11} + a_{12})b_{22}$$

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

$$m_6 = a_{22}(b_{21} - b_{11})$$

$$m_7 = (a_{21} + a_{22})b_{11}$$

$$c_{11} = m_1 + m_2 - m_4 + m_6$$

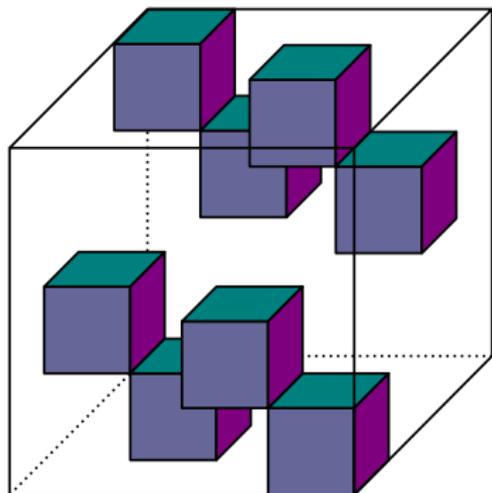
$$c_{12} = m_4 + m_5$$

$$c_{21} = m_6 + m_7$$

$$c_{22} = m_2 - m_3 + m_5 - m_7$$

7 moltiplicazioni e 18 addizioni (invece di 8+4) [Strassen, 1969]

## Il tensore del prodotto di matrici



$$\begin{aligned}\langle 2, 2, 2 \rangle &= (a_{11}b_{11} + a_{12}b_{21})c_{11} + \\ &\quad (a_{11}b_{12} + a_{12}b_{22})c_{12} + \\ &\quad (a_{21}b_{11} + a_{22}b_{21})c_{21} + \\ &\quad (a_{21}b_{12} + a_{22}b_{22})c_{22} \\ &= \sum_{i,j,k=1}^2 a_{ij}b_{jk}c_{ik}\end{aligned}$$

## Il rango di bordo

$$T := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad T \text{ ha rango } 3. \text{ Però...}$$

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & \varepsilon \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} \varepsilon^{-1} \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ \varepsilon \end{pmatrix} + \begin{pmatrix} 1 \\ \varepsilon^{-1} \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

o anche

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ \varepsilon \end{pmatrix} \otimes \begin{pmatrix} 1 \\ \varepsilon \end{pmatrix} + \begin{pmatrix} \varepsilon \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \varepsilon T + O(\varepsilon^2)$$

**Rango di bordo** di un tensore  $t \in k^{m \times n \times p}$ :

è il minimo  $r$  per cui esiste  $\tilde{t} \in k[\varepsilon]^{m \times n \times p}$  di rango  $r$   
con **termine di grado più basso** uguale a  $t$

Per il tensore  $t$  qui sopra,  $R(t) = 3$  ma  $\underline{R}(t) = 2$ .

## L'esempio di Bini

$$F = \text{tensore di } \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & 0 \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}$$

$$(a_{12} + \varepsilon a_{11})(b_{12} + \varepsilon b_{22})c_{21} +$$

$$(a_{21} + \varepsilon a_{11})b_{11}(c_{11} + \varepsilon c_{12}) -$$

$$a_{12}b_{12}(c_{11} + c_{21} + \varepsilon c_{22}) -$$

$$a_{21}(b_{11} + b_{12} + \varepsilon b_{21})c_{11} +$$

$$(a_{12} + a_{21})(b_{12} + \varepsilon b_{21})(c_{11} + \varepsilon c_{22}) = \varepsilon F + O(\varepsilon^2)$$

$\underline{R}(F) = 5$ . Usando questo risultato,  $\omega < 2.77989$  [Bini *et al.*, 1979]

# Il $\tau$ -teorema di Schönhage

Teorema (Schönhage, 1981)

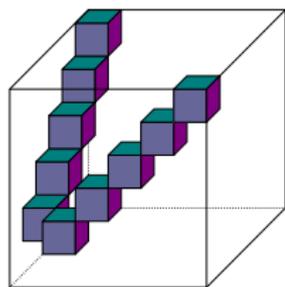
$$\underline{R}\left(\bigoplus_{i=1}^s \langle m_i, n_i, p_i \rangle\right) \leq r \implies \sum_{i=1}^s (m_i n_i p_i)^{\omega/3} \leq r$$

Esempio di Schönhage

$$\underline{R}(\langle h, 1, k \rangle \oplus \langle 1, (h-1)(k-1), 1 \rangle) \leq hk + 1$$

- ▶ Il *border rank* non è additivo
- ▶  $h = k = 4 \implies \omega \leq 2.55$  [Schönhage, 1981]

## Il metodo laser di Strassen



$$\underline{R} \left( \sum_{i=1}^5 a_i b_0 c_i + a_0 b_i c_i \right) \leq 6$$

Componiamone (con  $\otimes$ ) tre copie:

$$\sum_{i,j,k=1\dots 5} \left( a_{ijo} b_{0jk} c_{i0k} + a_{ijk} b_{0jk} c_{i00} + a_{ijo} b_{00k} c_{ijk} + a_{ijk} b_{00k} c_{ij0} + \right. \\ \left. a_{0jo} b_{ijk} c_{i0k} + a_{0jk} b_{ijk} c_{i00} + a_{0jo} b_{i0k} c_{ijk} + a_{0jk} b_{i0k} c_{ij0} \right)$$

► ha la struttura “a blocchi” di un prodotto  $\langle 2, 2, 2 \rangle$ :

$$\begin{pmatrix} a_{ijo} & a_{0jo} \\ a_{ijk} & a_{0jk} \end{pmatrix} \begin{pmatrix} b_{0jk} & b_{00k} \\ b_{ijk} & b_{i0k} \end{pmatrix} = \begin{pmatrix} c_{i0k} & c_{ijk} \\ c_{i00} & c_{ij0} \end{pmatrix}$$

► il prodotto in ogni blocco è di tipo  $\langle m, n, p \rangle$  con  $mnp = 5^3$

Utilizzando questo tensore,  $\omega \leq 2.479$  [Strassen, 1987]

# La costruzione di Coppersmith e Winograd

- ▶ Struttura a blocchi:  $N$ -esima potenza di un  **tensore planare**
- ▶ Famiglia di funzioni  $A_w, B_w, C_w$  degli indici delle  $a, b, c$  tali che  $A_w + B_w + C_w = 0 \quad \forall w$
- ▶ Se una di queste funzioni fosse iniettiva, sapremmo estrarre una  **diagonale**
- ▶  **Mediamo**  su tutte le possibili scelte di  $w$ : la funzione “media” è quasi iniettiva e fornisce una diagonale “abbastanza grande”
- ▶  $\tau$ -teorema sulla diagonale trovata:  $\omega < 2.41$
- ▶ Introducendo dei “pesi” per gli elementi della diagonale,  $\omega < 2.39$
- ▶ Raffinando queste idee, si arriva a  $\omega < 2.376$  [Coppersmith e Winograd, 1987]