

Matrix Multiplication (MM)

- Given two matrices $x = (x_{ij})_{i,j \in [n]}$ &
 $y = (y_{ij})_{i,j \in [n]}$ over R .

We want to compute $x \cdot y =: z = (z_{ij})_{n \times n}$

$$\triangleright z_{ij} = \sum_{k \in [n]} x_{ik} \cdot y_{kj}$$

\triangleright Naively, MM requires n^3 R -mult. &
 $n^2(n-1)$ R -addn.

- Qn: Could we reduce R-mult. at the cost of R-addn. ? (Say, n fixed ?)

▷ Strassen (1969) showed how to multiply 2×2 matrices using $2^3 - 1 = \underline{7}$ mult. (but 18 addn.).

The 7 products:

$$p_1 = (x_{11} + x_{22})(y_{11} + y_{22})$$

$$p_2 = (x_{21} + x_{22})y_{11}$$

$$p_3 = x_{11}(y_{12} - y_{22})$$

$$p_7 = (x_{12} - x_{22})(y_{21} + y_{22})$$

|
|
|
|

$$p_4 = x_{22}(-y_{11} + y_{21})$$

$$p_5 = (x_{11} + x_{12})y_{22}$$

$$p_6 = (-x_{11} + x_{21})(y_{11} + y_{12})$$

$$\triangleright x \cdot y = z = \begin{pmatrix} p_1 + p_4 - p_5 + p_7 & p_3 + p_5 \\ p_2 + p_4 & p_1 + p_3 - p_2 + p_6 \end{pmatrix}$$

- Since, the above holds for any ring R ,
we can apply this to design a recursive
algorithm for MM. [Use halving of n .]

Theorem [Strassen '69]: MM takes $O(n^{\lg 7})$ R -ops.



Pen



Eraser



Text



Undo



Redo

SAVE

MORE

$$\triangleright x \cdot y = z = \begin{pmatrix} p_1 + p_4 - p_5 + p_7 & p_3 + p_5 \\ p_2 + p_4 & p_1 + p_3 - p_2 + p_6 \end{pmatrix}$$

- Since, the above holds for any ring R , we can apply this to design a recursive algorithm for MM. [Use halving of n .]

Theorem [Strassen '69]: MM takes $O(n^{\lg 7})$ R -ops.

Pf:

- Let $x, y \in R^{n \times n}$, $n =: 2^l$.
- We'll show by induction on l that we can do MM in

7^l R-mult. & $6 \cdot (7^l - 4^l)$ R-addns.

Base case: $[l=1]$: already seen.

Ind. step $[l-1 \rightarrow l]$: Use block-structure mult.:

$$\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} \cdot \begin{pmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{pmatrix} =: \begin{pmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{pmatrix}_{2^l \times 2^l}$$

Where, x_{ij}, y_{ij}, z_{ij} 's are $2^{l-1} \times 2^{l-1}$ matrices.

So, 1) Use Strassen's eqns. of 2×2 matrices (over $R^{\frac{n}{2} \times \frac{n}{2}}$)

2) & recursion to $\frac{n}{2}$.

Time taken: (by induction)

$$\#R\text{-mult} = 7 \times (7^{l-1}) = 7^l \quad \checkmark$$

$$\#R\text{-addns} = 7 \times 6(7^{l-1} - 4^{l-1}) + 18 \times (2^{l-1})^2$$

7 recursive calls

Strassen's eqns

$$= 6 \cdot (7^l - 4^l) \quad \checkmark$$

\Rightarrow Overall, $O(7^l) = O(n^{2.8})$ R-ops. \square

- After decades of work, the current best for MM is $O(n^{2.3728639})$ [Le Gall, 2014]

Conjecture: MM has complexity $O(n^{2+\epsilon})$, for any $\epsilon > 0$

The exponent of MM

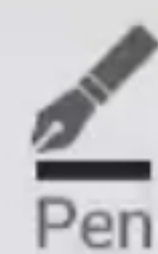
- Let's denote it by ω .

$$\triangleright 2 \leq \omega < 2.3728639.$$

- All the known upper bounds on ω use the notion of tensor rank.

Defn: The MM-tensor is a polynomial in $R[X_{ij}, Y_{ij}, Z_{ij} \mid i \leq j \in [n]]$ namely:

$$\underline{T_{n,R}} := \sum_{i,j,k \in [n]} X_{ik} Y_{kj} Z_{ij} \quad (\text{cubic poly.})$$



Pen



Eraser



Text



Undo



Redo

SAVE MORE

$$\text{eg. } T_{Z,R} = Z_{11}(X_{11}Y_{11} + X_{12}Y_{21}) + \\ Z_{12}(X_{11}Y_{12} + X_{12}Y_{22}) + Z_{21}(X_{21}Y_{11} + X_{22}Y_{21}) \\ + Z_{22}(X_{21}Y_{12} + X_{22}Y_{22}).$$

Defn: Rank $r(T)$ of tensor T is the least r st. \exists linear forms $L_i \in R[\bar{x}]$, $M_i \in R[\bar{y}]$, $N_i \in R[\bar{z}]$, $i \in [r]$ satisfying:

$$T = \sum_{i \in [r]} L_i \cdot M_i \cdot N_i$$

simple tensor of $rk=1$

\rightarrow order-2 tensor: $T(\bar{x}, \bar{y}) = \sum_{i \in [r]} L_i(\bar{x}) \cdot M_i(\bar{y})$
[Exercise: Matrix Rank?]





Pen



Eraser



Text



Undo



Redo

SAVE

MORE

$$\triangleright n^2 \leq r(T_{n,R}) \leq n^3.$$

Pf sketch: [$\leq n^3$]: by defn of r_k .

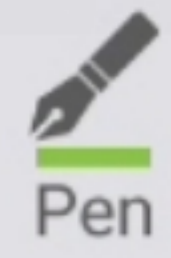
[$\geq n^2$]: Fixing variables in $T_{n,R}$ gives a contradiction. \square

Claim: MM_n can be done in $r(T_{n,R})$ many R -multiplications.

Pf: • Say, $T_{n,R} = \sum_{i \in [r]} L_i(\bar{x}) \cdot M_i(\bar{y}) \cdot N_i(\bar{z})$

- Extract the coefficient of z_{11} both sides & so on.
- Products $\{L_i \cdot M_i \mid i \in [r]\}$ are enough. \square





Pen



Eraser



Text



Undo



Redo

SAVE

MORE

— Strassen's eqns show $r_2(T_{2,R}) \leq 7$.

OPEN: $r_2(T_{3,R})$ is unknown.
▷ " $\in [19, 23]$.

[Håstad '90]: (order-3) tensor rank computation is NP-hard.

▷ $r_2(T_{n_0}) = \dots = r_0 \implies \omega \leq \log_{n_0} r_0$.

→ Known results define tensors "related to" $T_{n,R}$.

