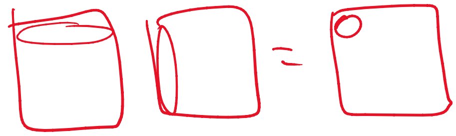
Matrix Multiplication

Problem Given $n \times n$ matrices $A = (a_{ij})$, $B = (b_{ij})$

compute $C = AB$

where $c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$



trivial: $\boxed{O(n^3)}$ time

naive lower bd: $\Omega(n^2)$.

Strassen's Alg'm ('69)

Warm-up: $n = 2$

to compute

$c_{11} = a_{11} b_{11} + a_{12} b_{21}$ $\quad c_{12} = a_{11} b_{12} + a_{12} b_{22}$

→ $c_{21} = a_{21} b_{11} + a_{22} b_{21}$ $\quad c_{22} = a_{21} b_{12} + a_{22} b_{22}$

naively: 8 mults.

idea - power of subtraction, again!

$P_1 = a_{11}(b_{11} - b_{21})$ $\quad P_2 = (a_{11} + a_{12}) b_{21}$

$P_3 = a_{22}(b_{22} - b_{12})$ $\quad P_4 = (a_{21} + a_{22}) b_{12}$

$P_5 = (a_{11} + a_{22})(b_{21} + b_{12})$

$P_6 = (a_{12} - a_{22})(b_{21} + b_{22})$ $\Rightarrow$ 7 mults.

$P_7 = (a_{21} - a_{11})(b_{11} + b_{12})$ by magic!

$c_{11} = P_1 + P_2$ $\quad c_{22} = P_3 + P_4$

$c_{12} = P_5 + P_6 + P_3 - P_2$ $\quad c_{21} = P_5 + P_7 + P_1 - P_4$

General n: idea - D & C

$A = \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right)$ $\quad B = \left(\begin{array}{c|c} B_{11} & B_{12} \\ \hline B_{21} & B_{22} \end{array}\right)$

each A_{ij}, B_{ij} is $(n/2) \times (n/2)$ matrix

$C = \left(\begin{array}{c|c} C_{11} & C_{12} \end{array}\right.$

$$\Rightarrow \quad C = \left(\begin{array}{c|c} C_{11} & C_{12} \\ \hline C_{21} & C_{22} \end{array}\right)$$

Same formula

$$\Rightarrow \quad T(n) = 7\,T\left(\frac{n}{2}\right) + O(n^2)$$

$$\Rightarrow \quad T(n) = O(n^{\log_2 7}) \leq \boxed{O(n^{2.81})}.$$

Better?

History

3-way D&C like Toom & Cook!

$$T(n) = \underbrace{23}_{\text{Laderman '76}}\,T\left(\frac{n}{3}\right) + O(n^2)$$

$$\Rightarrow O(n^{\log_3 23}) \leq O(n^{2.85})$$

worse!

Pan '78:

$$T(n) = 143640\,T\left(\frac{n}{70}\right) + O(n^2)$$

(143640 = $\frac{k^3 - 4k}{3} + 6k^2$, with 70 ← k)

$$\Rightarrow O(n^{\log_{70} 143640}) \leq O(n^{2.796})$$

Pan '79: $T(n) = 41952\,T\left(\frac{n}{46}\right) + O(n^2) \Rightarrow O(n^{2.781})$

Bini et al. '80: $O(n^{2.780})$ ("border rank")

→ Schönhage '81: $O(n^{2.522})$

Strassen '86: $O(n^{2.479})$ ("laser method")

Coppersmith & Winograd '90: $O(n^{2.376})$ ←

Stothers '10: $O(n^{2.374})$

Williams '12: $O(n^{2.373})$

Stothers '10: $O(n^{\dots})$

Vassilevska-Williams '12: $O(n^{2.373})$

Le Gall '14: $O(n^{2.37287})$

Alman & Vassilevska-W. '21: $O(n^{2.37286})$

Major Open Problem:

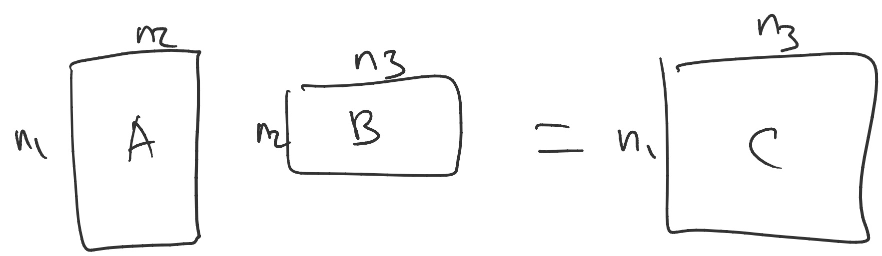
determine the <u>matrix multiplication exponent</u>

denoted by ω

$$2 \leq \omega < 2.37286$$

my notes - $n^{2.548}$

Rectangular Matrix Mult.



let $M(n_1, n_2, n_3)$ = time to multiply $n_1 \times n_2$ & $n_2 \times n_3$ matrices

$\omega(a,b,c)$ = exponent for $M(n^a, n^b, n^c)$

Known facts:

1. $\omega(\cdot,\cdot,\cdot)$ is convex

   $\left(\omega(ta_1+(1-t)a_2,\ tb_1+(1-t)b_2,\ tc_1+(1-t)c_2) \leq t\,\omega(a_1,b_1,c_1) + (1-t)\,\omega(a_2,b_2,c_2) \quad \forall t \in [0,1] \right)$

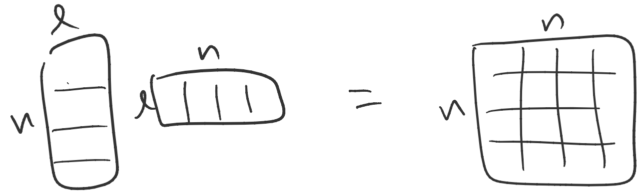
2. $\omega(\cdot,\cdot,\cdot)$ is symmetry

$$\left(\omega(a,b,c) = \omega(c,b,a) = \omega(b,c,a) = \dots \right)$$

Quick Bds:

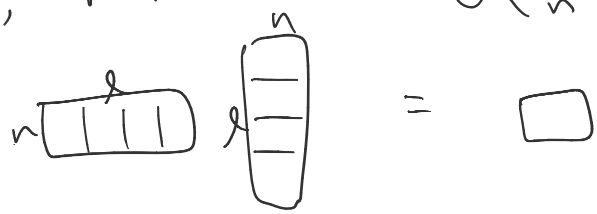
if $\ell \le n$, $\quad M(n,\ell,n) = O\left(\left(\frac{n}{\ell}\right)^2 \cdot \ell^{\omega}\right) = O(\ell^{\omega-2} n^2)$.

skinny case



if $\ell > n$, $\quad M(n,\ell,n) = O\left(\frac{\ell}{n} \cdot n^{\omega}\right) = O(\ell \cdot n^{\omega-1})$

fat case



but there are better bds . . .

e.g. skinny case:

Coppersmith '82: $\quad M(n, n^{0.172}, n) = \tilde{O}(n^2)$

⋮

→ Le Gall & Urrutia '18: $\quad M(n, n^{0.31389}, n) = O(n^{2+\varepsilon})$

↑ vect. matrix mult. exponent, denoted by α.

$$1 \ge \alpha \ge 0.31389$$

e.g. fat case:

$k > 1 \quad M(n, n^k, n) = O(n^{k+1+f(k)})$

where $f(k) \to 0$ as $k \to \infty$.

⋮