Matrix and Vectors

$A \in \mathbb{R}^{n \times m}$ ← A is n by m matrix whose entry is in $\mathbb{R}$

$$A = \underbrace{\left.\begin{bmatrix} A_{11} & \cdots & A_{1m} \\ \vdots & \ddots & \vdots \\ A_{n1} & \cdots & A_{nm} \end{bmatrix}\right\} n \text{ rows}}_{m \text{ columns}}$$

Vectors : $1 \times n$ or $n \times 1$ matrix

$a = [a_1 \quad \cdots \quad a_n]$ row vector

$\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$ column vector.

("usually" we'll use column vectors unless specified)

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1m} \\ \vdots & \ddots & \vdots \\ A_{n1} & \cdots & A_{nm} \end{bmatrix} = \begin{bmatrix} - & r_1 & - \\ & \vdots & \\ - & r_n & - \end{bmatrix} = \begin{bmatrix} | & & | \\ c_1 & \cdots & c_m \\ | & & | \end{bmatrix}$$

$\forall i \in [n]$: r_i is a row vector s.t. $\forall j \in [m]$, $(r_i)_j = A_{ij}$.

$\forall j \in [m]$: c_j is a column vector s.t. $\forall i \in [n]$, $(c_j)_i = A_{ij}$.

Matrix Multiplication: Suppose $A \in \mathbb{R}^{n \times m}$ and $B \in \mathbb{R}^{n' \times m'}$.

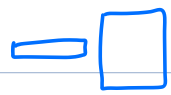
$A+B$: defined when $\begin{matrix} n=n' \\ m=m' \end{matrix}$ and $\forall (i,j) \in [n] \times [m]: (A+B)_{ij} = A_{ij} + B_{ij}$.

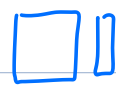
AB: defined when $m = n'$. $AB \in \mathbb{R}^{n \times m'}$ and

$$\forall (i,j) \in [n] \times [m'] : (AB)_{ij} = \sum_{k=1}^{m} A_{ik} \cdot B_{kj}.$$

Special cases:

① A is a row vector ($n=1$) or B is a column vector ($m'=1$)





② A is a row vector ($n=1$) **AND** B is a column vector ($m'=1$)

Then AB is just one number, which is $\underline{\sum_{i=1}^{m} A_i \cdot B_i}$.

inner/dot product of two vectors A, B. Will denote by $A \cdot B$ or $\langle A, B \rangle$.

(fun facts: if $u, v \in \mathbb{R}^n$,

then $\langle u, v \rangle = 0 \Leftrightarrow u, v$ are orthogonal.

$\|u\|_2^2 = \langle u, u \rangle =$ squared length of u)

Inner product interpretation:

$A \in \mathbb{R}^{n \times m}$, $B \in \mathbb{R}^{m \times m'}$

$$A = \begin{bmatrix} - & a_1 & - \\ & \vdots & \\ - & a_n & - \end{bmatrix}, \quad B = \begin{bmatrix} | & & | \\ b_1 & \cdots & b_{m'} \\ | & & | \end{bmatrix}, \quad \text{then } (AB)_{ij} = \langle a_i, b_j \rangle.$$

Strassen's ALG

Let $A, B \in \mathbb{R}^{n \times n}$. Want to compute $C = A \cdot B \in \mathbb{R}^{n \times n}$.

Naive: $O(n^3)$ time. Can do better?

Divide and Conquer: Split A into 4 "blocks" $A_{11}, A_{12}, A_{21}, A_{22}$

$$A = \begin{bmatrix} A^{11} & A^{12} \\ A^{21} & A^{22} \end{bmatrix}, \quad A^{ij} \in \mathbb{R}^{n/2 \times n/2}$$ (assume n even)

(e.g., $\forall i,j \in [n/2]$, $(A^{11})_{ij} = A_{ij}$, $(A^{22})_{ij} = A_{i+n/2, j+n/2}$)

Then, if we write $C = A \times B$

$$\begin{bmatrix} C^{11} & C^{12} \\ C^{21} & C^{22} \end{bmatrix} = \begin{bmatrix} A^{11} & A^{12} \\ A^{21} & A^{22} \end{bmatrix} \times \begin{bmatrix} B^{11} & B^{12} \\ B^{21} & B^{22} \end{bmatrix}$$

What can we say about C^{ij}, A^{ij}, B^{ij}s?

If $n=2$ and they were numbers, then $C^{ij} = A^{i1} B^{1j} + A^{i2} B^{2j}$.

__Lemma__ It is true for any (even) n.

__Pf,__ Let's just prove $C^{11} = A^{11} B^{11} + A^{12} B^{21}$ (others are similar)

$\forall i,j \in [n/2]$, $(C^{11})_{ij} = C_{ij} = \langle a_i, b_j \rangle$

$$A = \begin{bmatrix} — a_1 — \\ \vdots \\ — a_n — \end{bmatrix} = \begin{bmatrix} \overbrace{\begin{matrix} — a_1^1 — \\ \vdots \\ — a_{n/2}^1 — \end{matrix}}^{A^{11}} & \overbrace{\begin{matrix} — a_1^2 — \\ \\ — a_{n/2}^2 — \end{matrix}}^{A^{12}} \\ \underbrace{\begin{matrix} — a_{n/2+1}^1 — \\ \vdots \\ — a_n^1 — \end{matrix}}_{A^{21}} & \underbrace{\begin{matrix} — a_{n/2+1}^2 — \\ \\ — a_n^2 — \end{matrix}}_{A^{22}} \end{bmatrix}$$

$$B = \begin{bmatrix} | & & | \\ b_1 & \cdots & b_n \\ | & & | \end{bmatrix} = \begin{bmatrix} \overbrace{\begin{matrix} | \\ b_1^1 \\ | \end{matrix}}^{B^{11}} & \cdots & \overbrace{\begin{matrix} | \\ b_n^1 \\ | \end{matrix}}^{B^{12}} \\ \underbrace{\begin{matrix} | \\ b_1^2 \\ | \end{matrix}}_{B^{21}} & & \underbrace{\begin{matrix} | \\ b_n^2 \\ | \end{matrix}}_{B^{22}} \end{bmatrix}$$

Then, $\langle a_i, b_j \rangle = \sum_{k=1}^{n} (a_i)_k (b_j)_k = \sum_{k=1}^{n/2} (a_i)_k (b_j)_k + \sum_{k=n/2+1}^{n} (a_i)_k (b_j)_k$

$= \langle a_i^1, b_j^1 \rangle + \langle a_i^2, b_j^2 \rangle.$

$= (A^{11}B^{11})_{ij} + (A^{12}B^{21})_{ij}$ $\square$

So, $C^{11} = A^{11}B^{11} + A^{12}B^{21}$

$C^{12} = A^{11}B^{12} + A^{12}B^{22}$

$C^{21} = A^{21}B^{11} + A^{22}B^{21}$

$C^{22} = A^{21}B^{12} + A^{22}B^{22}$

$O(n^2)$

4 additions of $n/2 \times n/2$ matrices.

8 multiplications of $n/2 \times n/2$ " .

$8 \cdot T(n/2)$.

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

Master thm with $k=8$, $b=2$, $d=2 \Rightarrow T(n) = O(n^3)$.

But following "magic" will reduce 8 to 7.

$S^1 = B^{12} - B^{22}$, $S^2 = A^{11} + A^{12}$, $S^3 = A^{21} + A^{22}$, $S^4 = B^{21} - B^{11}$, $S^5 = A^{11} + A^{22}$

$S^6 = B^{11} + B^{22}$, $S^7 = A^{12} - A^{22}$, $S^8 = B^{21} + B^{22}$, $S^9 = A^{11} - A^{21}$, $S^{10} = B^{11} + B^{12}$

$P^1 = A^{11} S^1$, $P^2 = S^2 \cdot B^{22}$, $P^3 = S^3 \cdot B^{11}$, $P^4 = A^{22} \cdot S^4$, $P^5 = S^5 \cdot S^6$

$P^6 = S^7 \cdot S^8$, $S^7 = S^9 \cdot S^{10}$.

$C^{11} = P^5 + P^4 - P^2 + P^6$

$C^{12} = P^1 + P^2$

$C^{21} = P^3 + P^4$

$C^{22} = P^5 + P^1 - P^3 - P^7$

So, 18 additions but "7" multiplications.

Running time $T(n) = 7T(n/2) + O(n^2)$.

Master thm with $k=7$, $b=2$, $d=2$ $\Longrightarrow$ $T(n) = O(n^{\log_2 8}) = O(n^{2.73\cdots})$.

Current record [DWZ23]: $O(n^{2.37\cdots})$

Ran Duan, PhD Michigan 2011.