Matrix Multiplication

Note Title 10/30/2007

Matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & \ddots & & & \vdots \\ a_{31} & & \ddots & & \vdots \\ \vdots & & & \ddots & \vdots \\ a_{m1} & \cdots & \cdots & \cdots & a_{mn} \end{bmatrix}$$

a_{ij} refers to the element in the ith ROW and the jth COLUMN.

row number always goes first.

For today: Assume square matrices.

Note Title 10/29/2007

Faster Matrix Multiplication

We can define a faster form of matrix multiplication in terms of $\underbrace{MM}_{\text{Basic Matrix Multiplication}} + \underbrace{MA}_{\text{Basic Matrix Addition.}}$

Let's start with Matrix Addition.

Matrix Addition (MA)

Note Title 10/29/2007

If $C = A + B$, then A and B both are $m \times n$ matrices.

Let's take $m = n$

Adding two matrices:

$$\underset{A}{\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}} \underset{+}{+} \underset{B}{\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}} \underset{=}{=} \underset{C}{\begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} \\ a_{21}+b_{21} & a_{22}+b_{22} \end{pmatrix}}$$

Example:

$$\begin{pmatrix} 1 & 3 \\ 5 & 7 \end{pmatrix} + \begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix} = \begin{pmatrix} 3 & 7 \\ 11 & 15 \end{pmatrix}$$

In general:

$$\underset{A}{\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}} \underset{+}{+} \underset{B}{\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}} \underset{=}{=} \underset{C}{\begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}}$$

Where $c_{ij} = a_{ij} + b_{ij}$.

So if the matrix is $n \times n$, MA is an n^2 algorithm (one addition for each position in the C matrix).

$\Downarrow$

One addition times n^2 positions $= n^2$

So for 2×2, there are 4 additions.

Matrix Multiplication (MM)

If $C = A \times B$, then
$\underset{m \times n}{A}$, $\underset{n \times p}{B}$

Let's take $m = n = p$

Multiplying two matrices:

$$\underset{A}{\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}} \underset{\cdot}{\cdot} \underset{B}{\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}} \underset{=}{=} \underset{C}{\begin{pmatrix} (a_{11} \cdot b_{11} + a_{12} \cdot b_{21}) & (a_{11} \cdot b_{12} + a_{12} \cdot b_{22}) \\ (a_{21} \cdot b_{11} + a_{22} \cdot b_{21}) & (a_{21} \cdot b_{12} + a_{22} \cdot b_{22}) \end{pmatrix}}$$

Example:

$$\begin{pmatrix} 1 & 3 \\ 5 & 7 \end{pmatrix} \cdot \begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix} = \begin{pmatrix} 2+18 & 4+24 \\ 10+42 & 20+56 \end{pmatrix} = \begin{pmatrix} 20 & 28 \\ 52 & 76 \end{pmatrix}$$

In general:

$$\underset{A}{\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}} \underset{\cdot}{\cdot} \underset{B}{\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}} \underset{=}{=} \underset{C}{\begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}}$$

where $c_{ij} = \sum_{k=1}^{2} (a_{ik} \cdot b_{kj})$ $(n=2)$

So if the matrix is $n \times n$, MM is an n^3 algorithm (n multiplications for each position in the C matrix).

$\Downarrow$

n multiplications (and $n-1$ additions) times n^2 positions $= n^3$

So for 2×2, there are $4 \cdot (2M + 1A)$
$= 8$ multiplications $+ 4$ additions.

What if we had a pair of 8 x 8 matrices to multiply?

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{18} \\ a_{21} & & & \\ \vdots & & & \\ a_{81} & & & \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{18} \\ b_{21} & & & \\ \vdots & & & \\ b_{81} & & & \end{pmatrix}$$

Could we subdivide the problem?

$$\left(\begin{array}{ccc|ccc} a_{11} & \cdots & a_{14} & a_{15} & \cdots & a_{18} \\ \vdots & A_{11} & & & A_{12} & \\ a_{41} & & & & & \\ \hline a_{51} & & & & & \\ \vdots & A_{21} & & & A_{22} & \\ a_{81} & & & & & \end{array}\right) \left(\begin{array}{ccc|ccc} b_{11} & \cdots & b_{14} & b_{15} & \cdots & b_{18} \\ \vdots & B_{11} & & & B_{12} & \\ b_{41} & & & & & \\ \hline b_{51} & & & & & \\ \vdots & B_{21} & & & B_{22} & \\ b_{81} & & & & & \end{array}\right)$$

Let $A_{11} = \begin{pmatrix} a_{11} & \cdots & a_{14} \\ \vdots & & \vdots \\ a_{41} & & a_{44} \end{pmatrix}$ $A_{12} = \begin{pmatrix} a_{15} & \cdots & a_{18} \\ \vdots & & \\ a_{45} & \cdots & a_{48} \end{pmatrix}$

$B_{11} = \begin{pmatrix} b_{11} & \cdots & b_{14} \\ \vdots & & \\ b_{41} & \cdots & b_{44} \end{pmatrix}$

$B_{21} = \begin{pmatrix} b_{51} & \cdots & b_{54} \\ \vdots & & \vdots \\ b_{81} & & b_{84} \end{pmatrix}$

Note: $C = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix}$

$$\underbrace{C_{11}}_{\text{Matrix}} = \underbrace{A_{11} \cdot B_{11}}_{MM} \overbrace{+}^{MA} \underbrace{A_{12} \cdot B_{21}}_{MM}$$

So $\underset{\text{element}}{\underline{c_{11}}} = a_{11} \cdot b_{11} + a_{12} \cdot b_{21} + \ldots + a_{14} \cdot b_{41}$

$+$

$a_{15} \cdot b_{51} + \ldots \ldots + a_{18} \cdot b_{81}$

Which is what we wanted.

Total is $4 \times (2\,MM + 1\,MA)$

So what is the <u>overall</u> runtime?

Using brute force algorithms that we saw earlier, we get runtimes of:

$$MM(n) = O(n^3)$$
$$MA(n) = O(n^2)$$

If we use divide + conquer approach above, what do we get?

$$(n \times n) \cdot (n \times n)$$

$$\Rightarrow \begin{pmatrix} (\frac{1}{2}n \times \frac{1}{2}n) & (\quad) \\ (\quad) & (\quad) \end{pmatrix} \begin{pmatrix} (\frac{1}{2}n \times \frac{1}{2}n) & (\quad) \\ (\quad) & (\quad) \end{pmatrix}$$

So...

$$MM(n) = 4\left[2MM\left(\tfrac{n}{2}\right) + MA\left(\tfrac{n}{2}\right)\right]$$

$$= 8MM\left(\tfrac{n}{2}\right) + 4MA\left(\tfrac{n}{2}\right)$$

We go all the way down to the base case: $MM(2) = 8M + 4A$.

The one place where we do basic multiplication + addition.

Let's assume MA can't be done in better than $O(n^2)$ time.

$4\left(\frac{n^2}{2^2}\right) = O(n^2)$

$$MM(n) = 8MM\left(\tfrac{n}{2}\right) + O(n^2)$$

We can apply Master Theorem.

$a = 8$

$b = 2$

$f(n) = n^2$

$n^{\log_2 8} = n^3$

Which case applies?

$$f(n) = n^2 \in O(n^3)$$

Case I of M.T.

$$\text{So } MM(n) = \Theta(n^3)$$

What if we could do 2x2 MM better??

Theorem

If we had a way to multiply two 2×2 matrices using only 7 multiplication operations, and a constant number of additional addition operations, then we could improve the runtime of the general matrix multiplication problem from $O(n^3)$.

$$\Downarrow$$

Proof:

Assuming two 2×2 matrices can be multiplied using 7 multiplication operations and y addition operations:

$$MM(n) = 7 \cdot MM\left(\frac{n}{2}\right) + y \cdot MA\left(\frac{n}{2}\right)$$
$$= 7 \cdot MM\left(\frac{n}{2}\right) + O(n^2)$$

We can apply Master Theorem.

$a = 7$ $\quad n^{\log_2 7} = n^{2.-}$
$b = 2$
$f(n) = n^2$

$\Downarrow$ Which case applies?

$f(n) = n^2 \in O(n^{\log_2 7})$ $\quad o(n^{3-\alpha})$

Case I of MT $\quad$ So $MM(n) = \Theta(n^{\log_2 7})$

which is better than n^3 runtime.

Thought Questions

1. Is there a way that we can multiply two 2x2 matrices with only 7 multiplication operations and a constant number of addition operations?

2. Do you think that there is any way we can do matrix multiplication in better than n^2 time?

What if we could multiply a 2x2 matrix in 7 multiplications??

Strassen

$$C = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix}$$

7 products

$$P_1 = (a_{11} + a_{22})(b_{11} + b_{22})$$

$$P_2 = (a_{21} + a_{22})b_{11}$$

$$P_3 = (a_{11})(b_{12} - b_{22})$$

$$P_4 = (a_{22})(b_{21} - b_{11})$$

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

$$P_6 = (a_{21} - a_{11})(b_{11} + b_{12})$$

$$P_7 = (a_{12} - a_{22})(b_{21} + b_{22})$$

$$C_{11} = P_1 + P_4 - P_5 + P_7$$

$$C_{12} = P_3 + P_5$$

$$C_{21} = P_2 + P_4$$

$$C_{22} = P_1 - P_2 + P_3 + P_6$$

so 7 multiplications

18 additions or subtractions

So $T(n) = 7T\left(\frac{n}{2}\right) + 18\left(\Theta(n^2)\right)$

time for matrix addition

Lower Bound for Matrix Multiplication

If $C = AB$, where A and B are $n \times n$ matrices

Then C has n^2 elements.

So lower bound is at least n^2.