Lec12-karatsuba-strassen

Saturday, September 21, 2024 8:26 PM

Example of divide & conquer for numerics

(Anatoly, 1962)

Karatsuba's alg for integer multiplication

```
  123
 x 321
-----
  123
 246
369
-----
39483
```

```
  10101110
x 01011101
----------
  10101110
 00000000
 10101110
 10101110
 10101110
 00000000
 10101110
----------
1111110011010 10
```

n = length of numbers

n numbers of n bits

$O(n^2)$-time

Break items into half-sized blocks:

Say x is n bits. Let $m = \frac{n}{2}$.

$$x = x_1 2^m + x_0$$

$$y = y_1 2^m + y_0$$

Then $xy = (x_1 2^m + x_0)(y_1 2^m + y_0) = x_1 y_1 2^{2m} + (x_1 y_0 + x_0 y_1) 2^m + x_0 y_0$

$O(m)$ bitshift; $O(m)$ bitshift

4 mults of size m bitstrings

But, $x_1 y_0 + x_0 y_1 = (x_1 + x_0)(y_1 + y_0) - x_1 y_1 - x_0 y_0$

one mult

Only 3 mults needed!

Let $p_0 = x_0 y_0$

$p_1 = x_1 y_1$

$p_2 = (x_1 + x_0)(y_1 + y_0) - p_0 - p_1$

Then $xy = p_1 2^{2m} + p_2 2^m + p_0$

linear work at each level of divide & conquer

Analysis: Let $n = 2^k$ for some k

linear work at each level of divide & conquer

Analysis: Let $n = 2^k$ for some k.

$$\frac{T(2^k)}{3^k} = \frac{3\,T(2^{k-1})}{3^k} + \frac{c\,2^k}{3^k} \quad \Bigg\} \quad \frac{T(2^k)}{3^k} = \frac{T(2^{k-1})}{3^{k-1}} + c\cdot\left(\frac{2}{3}\right)^k$$

By recursion,

$$\frac{T(2^k)}{3^k} = \frac{T(2^{k-1})}{3^{k-1}} + c\cdot\left(\frac{2}{3}\right)^k$$

$$= \frac{T(2^{k-2})}{3^{k-2}} + c\cdot\left(\frac{2}{3}\right)^{k-1} + c\cdot\left(\frac{2}{3}\right)^k$$

$$= \cdots = \underbrace{\frac{T(1)}{1}}_{\text{base case}} + \underbrace{c\sum_{j=1}^{k}\left(\frac{2}{3}\right)^j}_{\text{geon series}} \le \beta \quad \text{for some constant } \beta$$

$$\Rightarrow \frac{T(2^k)}{3^k} \le \beta$$

$$\Rightarrow T(2^k) \le \beta 3^k = \beta 2^{\log_2 3^k} = \beta 2^{k\log_2 3}$$

$$T(n) \le \beta\left(2^k\right)^{\log_2 3} = \beta n^{1.58\cdots} = O(n^{1.58\cdots}) \le O(n^2)$$

Matrix multiplication:

$$\begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix}\begin{bmatrix}5 & 6\\ 7 & 8\end{bmatrix} = \begin{bmatrix}5+2\cdot 7 & 6+2\cdot 8\\ 5\cdot 3+4\cdot 7 & 3\cdot 6+4\cdot 8\end{bmatrix} = \begin{bmatrix}19 & 22\\ 43 & 50\end{bmatrix}$$

row i $\begin{bmatrix} a_{i1} & \cdots & a_{in}\end{bmatrix}$ $\begin{bmatrix} b_{1j}\\ \vdots\\ b_{nj}\end{bmatrix}$ col j $= \begin{bmatrix} \sum_{k=1}^{n} a_{ik}b_{kj}\end{bmatrix}$

each entry takes $\Theta(n)$ to compute

n^2 entries $\Rightarrow \Theta(n^3)$

$n\times n$ $\quad$ $n\times n$ $\quad$ $n\times n$

$(\Theta(n^2))$

Notice: Matrix addition is much cheaper than multiplication

$(\Theta(n^2))$

Notice: Matrix addition is much cheaper than multiplication

Let's try divide (conquering for later)

Assume matrix is of size $2^m \times 2^m$ for some m (so we can break the matrix into quarters)

$n \times n$

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \cdot \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}$$

$A \in \mathbb{R}^{n \times n}$ $B \in \mathbb{R}^{n \times n}$ $C \in \mathbb{R}^{n \times n}$

$$C_{11} = A_{11}B_{11} + A_{12}B_{21} \qquad C_{12} = A_{11}B_{12} + A_{12}B_{22}$$

$$C_{21} = A_{21}B_{11} + A_{22}B_{21} \qquad C_{22} = A_{21}B_{12} + A_{22}B_{22}$$

8 half-size matrix multiplications

$(2^3 = 8)$

Intuition: $(X+Y)(Z+W) = XZ + YZ + XW + YW$

$(X+Y)(Z+W)$: 1 mult; $XZ + YZ + XW + YW$: 4 mults

Can we rewrite the formulae for $C_{11}, C_{12}, C_{21}, C_{22}$ to use fewer mults

Strassen:

$$P_1 = (A_{11} + A_{22})(B_{11} + B_{22})$$

$$P_2 = (A_{21} + A_{22})B_{11}$$

$$P_3 = A_{11}(B_{12} - B_{22})$$

$$P_4 = A_{22}(B_{21} - B_{11})$$

$$P_5 = (A_{11} + A_{12})B_{22}$$

$$P_6 = (A_{21} - A_{11})(B_{11} + B_{22})$$

$$P_7 = (A_{12} - A_{22})(B_{21} + B_{22})$$

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

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

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

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

Only 7 mults instead of 8

Apply idea recursively to $P_1, P_2, \ldots, P_7$ which are square matrices half the size

recursion

Apply idea recursively to $P_1, P_2, \ldots, P_7$ which are square matrices half the size

Recurrence: $T(n) = \overbrace{7\,T\left(\frac{n}{2}\right)}^{\text{recursion}} + \overbrace{cn^2}^{\text{additions}}$

$n = 2^m$: $\quad T(2^m) = 7\,T(2^{m-1}) + c\,4^m$

$\qquad\qquad \underbrace{}_{7^m} \quad \underbrace{\phantom{7\,T(2^{m-1})}}_{7^m} \quad \underbrace{}_{7^m}$

$$\Rightarrow \quad \frac{T(2^m)}{7^m} = \frac{T(2^{m-1})}{7^{m-1}} + c\left(\frac{4}{7}\right)^m$$

Recursive expansion:

$$= \sum_{k=1}^{m} c\left(\frac{4}{7}\right)^k + \frac{T(1)}{1}$$

$$\leq c\sum_{k=1}^{\infty}\left(\frac{4}{7}\right)^k + \frac{T(1)}{1} = c\cdot\frac{\frac{4}{7}}{1-\frac{4}{7}} + T(1) = \frac{4}{3}c + T(1) = \alpha \quad \text{for some constant } \alpha.$$

$$\Rightarrow \frac{T(2^m)}{7^m} \leq \alpha \quad \Rightarrow \quad T(2^m) \leq \alpha 7^m = \alpha 2^{m\log_2 7}$$

$$\Rightarrow T(n) \leq \alpha n^{\log_2 7} = \alpha n^{2.807\ldots} = O(n^{2.807\ldots})$$

Strassen first to show improvement over $O(n^3)$

Can we do better?

Volker Strassen, 1969, $O(n^{2.807})$

⋮

Coppersmith & Winograd, 1990, $O(n^{2.3754})$

Andrew Stothers, 2010, $O(n^{2.3736})$

Virginia Williams, 2013, $O(n^{2.3729})$ (CMU, PhD, 2008)

⋮

Williams, Xu, Xu, Zhao, 2023, $O(n^{2.371552})$

Master Theorem (Simplified)

Given a recurrence of form

$$T(n) = aT\left(\frac{n}{b}\right) + \Theta(n^i), \quad \text{constants } a \geq 1, \; b > 1,$$

Case 1: $i < \log_b a$ then $T(n) = \Theta\left(n^{\log_b a}\right)$ — work increasing as we go down the tree
($a > b^i$)

Case 2: $i = \log_b a$ then $T(n) = \Theta\left(n^i \log n\right)$ — work same at each level
($a = b^i$)

Case 3: $i > \log_b a$ then $T(n) = \Theta\left(n^i\right)$ — work decreasing as we go down the tree
($a < b^i$)

Ex. $T(n) = 2T\left(\frac{n}{2}\right) + cn. \Rightarrow 1$