Chained Matrix Multiplication

Introduction

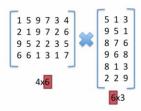
- · Matrix chain multiplication is an optimization problem that can be solved using dynamic progg.
- Given a sequence of matrices, we want to find the most efficient way to multiply these
 matrices together. The problem is not actually to perform the multiplications, but merely to
 decide in what order to perform the multiplications.
- Matrix multiplication is associative. No matter how we parenthesize the product, the final result will be the same.
- For example, if we have 3 matrices A1, A2, and A3, then

$${\bf A}_1{\bf A}_2{\bf A}_3=({\bf A}_1{\bf A}_2){\bf A}_3={\bf A}_1({\bf A}_2{\bf A}_3)$$

 Go through each possible parenthesization, this would require time O(2ⁿ), which is very slow and impractical for large n.

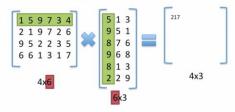
Matrix Chain Multiplication

Dot Product



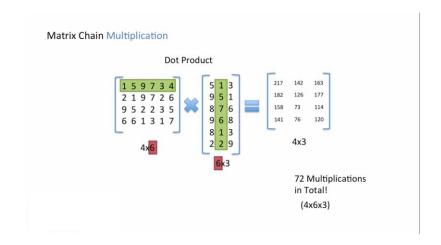
Matrix Chain Multiplication

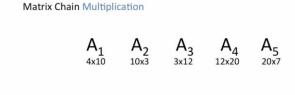
Dot Product

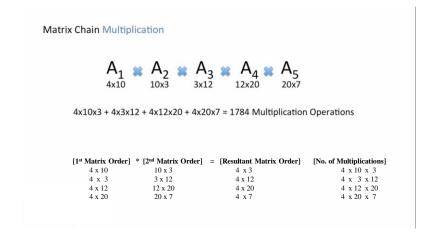


1x5 + 5x9 + 9x8 + 7x9 + 3x8 + 4x25 + 45 + 72 + 63 + 24 + 8 = 217

Dot Product | 1 5 9 7 3 4 | 2 1 9 7 2 6 | 9 5 2 2 3 5 6 6 1 3 1 7 | 4x6 | 5x3 | 4x3 | 5x3 | 5x3







Matrix Chain Multiplication

 $A_1 * A_2 * A_3 * A_4 * A_5$

Goal: Find the optimal way to multiply these matrices to perform the fewest multiplications.

Naïve Approach: Try them all, and pick the most optimal one.

Matrix Chain Multiplication

There is a better way! Dynamic Programming!
Step 1: Check if the problem has Optimal Substructure

Principle of Optimality

If we have an optimal solution for A_{i...j}

Assume the solution has the following parentheses:

 $(A_{i...k})(A_{k+1...j})$

If there is a better way to multiply $(A_{i...k})$, then we would have a more optimal solution. This would be a contradiction, as we already stated that we have the optimal solution for $A_{i...j}$ Therefore this problem has optimal substructure.

Where *k* is the *splitting point* that divides the problem into two halves and promises to provide the optimal solution.

Matrix Chain Multiplication

A matrix series $A_{i...j}$ can be broken up into a more efficient solution:

$$(A_{i...k})(A_{k+1...j})$$

We want to find out at which 'k' returns the fewest number of multiplications

We need to define our <u>recursive formula:</u> M[i,j] is the cost of multiplying matrices from A_i to A_i

Matrix Chain Multiplication

Now we want to try out a bunch of values for 'k' in order to see what the best one is:

$$M[i,j] = M[i,k] + M[k+1,j] + p_{i-1}p_kp_j$$

100 200 2x3x4

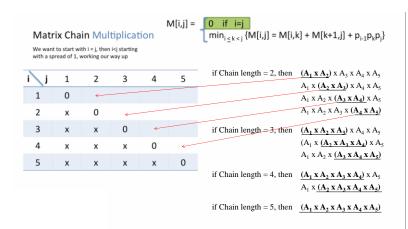
Since we don't know what k is, we try this range of k:

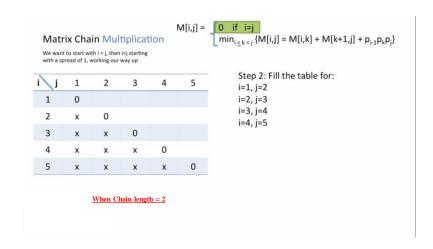
The minimum returned value is our solution!

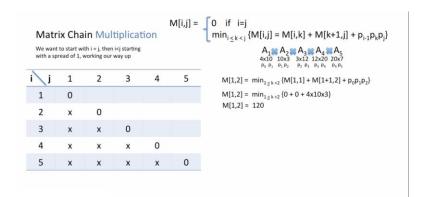
 $i \le k < j$

Our Final Recursive Formula:

$$M[i,j] = \begin{cases} 0 & \text{if } i=j \\ \min_{i \le k < j} \{M[i,j] = M[i,k] + M[k+1,j] + p_{i+1}p_kp_i\} \end{cases}$$







		ith i = j, then working our				$ \begin{bmatrix} 0 & \text{if } i = j \\ min_{i \le k < j} \left\{ M[i,j] = M[i,k] + M[k+1,j] + p_{i-1}p_kp_j \right\} \\ A_1 \not \cong A_2 \not \cong A_3 \not \cong A_4 \not \cong A_5 \\ 4 \times 10 & 10 \times 3 & 3 \times 12 & 12 \times 20 & 20 \times 7 \\ p_0 & p_1 & p_2 & p_2 & p_3 & p_4 & p_5 p_4 \end{aligned} $
\ j	1	2	3	4	5	$M[1,2] = \min_{1 \le k < 2} \{M[1,1] + M[1+1,2] + p_0p_1p_2\}$
1	0	120				$M[1,2] = \min_{1 \le k < 2} \{0 + 0 + 4x10x3\}$
2	x	0	360			M[1,2] = 120
3	х	х	0			$M[2,3] = min_{2 \leq k < 3} \{ M[2,2] + M[2+1,3] + p_1 p_2 p_3 \}$
4	х	×	x	0		$M[2,3] = \min_{2 \le k < 3} \{0 + 0 + 10x3x12\}$ M[2,3] = 360
5	х	х	х	х	0	
						$M[3,4] = \min_{3 \le k \le 4} \{M[3,3] + M[3+1,4] + p_2p_3p_4\}$ $M[3,4] = \min_{3 \le k \le 4} \{0 + 0 + 3x12x20\}$ $\frac{M[2,3]}{M[3,4]} = 720$ $M[3,4]$

$M[i,j] = \begin{bmatrix} 0 & \text{if } i=j \end{bmatrix}$ Matrix Chain Multiplication

We want to start with i = j, then i<j starting with a spread of 1, working our way up

i\j	1	2	3	4	5
1	0	120			
2	х	0	360		
3	х	х	0	720	
4	x	×	x	0	1680
5	х	х	х	х	0

$$\begin{cases} 0 & \text{if } i=j \\ \min_{i \le k < j} \{M[i,j] = M[i,k] + M[k+1,j] + p_{i-1}p_kp_j \} \end{cases}$$

 $M[4,5] = min_{4 \le k < 5} \{M[4,4] + M[4+1,5] + p_3p_4p_5\}$ $M[4,5] = \min_{4 \le k < 5} \{0 + 0 + 12x20x7\}$

 $\frac{M[1,2]}{M[4,5]} = 1680$ M [4,5]

		vith i = j, then working our				$\begin{array}{cccccccccccccccccccccccccccccccccccc$
\ j	1	2	3	4	5	$M[1,3] = \min_{1 \le k < 3}$
1	0	120				$\frac{k=1}{= M[1,1] + M[1+1,3] + p_0p_1p_3}$
2	x	0	360			= 0 + 360 + 4x10x12
3	х	х	0	720		= 840
4	x	×	×	0	1680	$\frac{k=2}{= M[1,2] + M[2+1,3] + p_0p_2p_3}$
5	x	х	х	х	0	= 120 + 0 + 4x3x12 = 264 <

Matrix Chain Multiplication

We want to start with i = j, then i<j starting with a spread of 1, working our way up

i\j	1	2	3	4	5
1	0	120	264		
2	x	0	360		
3	х	x	0	720	
4	х	x	х	0	1680
5	х	х	х	х	0

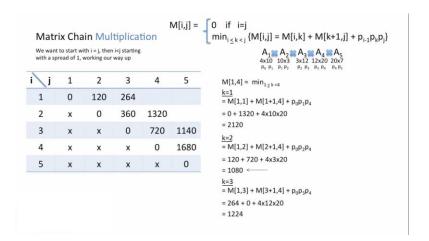
 $M[i,j] = \begin{bmatrix} 0 & \text{if } i=j \end{bmatrix}$ $\min_{i < k < j} \{M[i,j] = M[i,k] + M[k+1,j] + p_{i-1}p_kp_i\}$

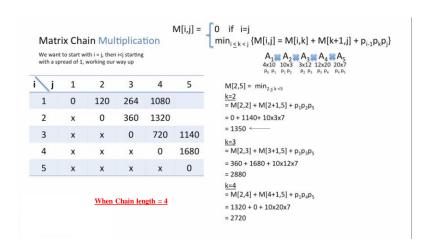
 $M[2,4] = \min_{2 \le k < 4}$ k=2 $= M[2,2] + M[2+1,4] + p_1p_2p_4$ = 0 + 720 + 10x3x20= 1320 < $= M[2,3] + M[3+1,4] + p_1p_3p_4$ $= 360 + 0 + 10 \times 12 \times 20$

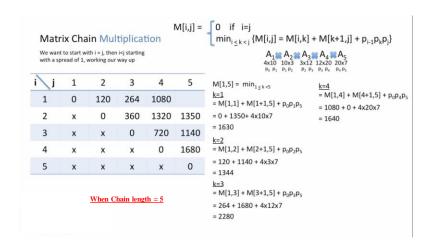
= 2760

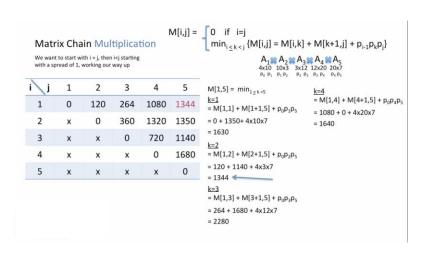
 $M[i,j] = \begin{bmatrix} 0 & \text{if } i=j \end{bmatrix}$ Matrix Chain Multiplication We want to start with i = j, then i<j starting with a spread of 1, working our way up 5 2 3 4 k=3 0 120 264 0 360 1320 = 1932 720 0 1680 0 X = 1140 <----

 $\min_{i < k < j} \{M[i,j] = M[i,k] + M[k+1,j] + p_{i-1}p_kp_i\}$ $M[3,5] = min_{3 \le k < 5}$ $= M[3,3] + M[3+1,5] + p_2p_3p_5$ = 0 + 1680 + 3x12x7 $= M[3,4] + M[4+1,5] + p_2p_4p_5$ = 720 + 0 + 3x20x7









$$\begin{split} M[i,j] = & \begin{bmatrix} 0 & \text{if } i = j \\ min_{i \leq k < j} \left\{ M[i,j] = M[i,k] + M[k+1,j] + p_{i+1}p_kp_j \right\} \\ A_1 & A_2 & A_3 & A_4 & A_5 \\ & & 4x10 & 10x3 & 3x12 & 12x20 & 20x7 \\ p_1, p_1, p_2, p_2, p_3, p_4, p_4, p_5, p_6 \end{aligned}$$

We now know that we can multiply A₁ to A₅ in as few as 1344 multiplication operations!

$$\begin{split} M[i,j] = & \int 0 & \text{if } i = j \\ min_{i \leq k < j} \left\{ M[i,j] = M[i,k] + M[k+1,j] + p_{j+1}p_kp_j \right\} \\ & A_1 \otimes A_2 \otimes A_3 \otimes A_4 \otimes A_5 \\ & A_1 \otimes A_2 \otimes A_3 \otimes A_4 \otimes A_5 \\ & A_1 \otimes A_2 \otimes A_3 \otimes A_4 \otimes A_5 \\ & A_2 \otimes A_3 \otimes A_4 \otimes A_5 \\ & A_3 \otimes A_4 \otimes A_5 \\ & A_4 \otimes A_5 \otimes A_5 \otimes A_6 \otimes A_5 \\ & A_4 \otimes A_5 \otimes A_5 \otimes A_6 \otimes A_5 \\ & A_5 \otimes A_5 \otimes A_6 \otimes A_5 \otimes A_6 \otimes A_5 \\ & A_5 \otimes A_5 \otimes A_6 \otimes A_6 \otimes A_5 \otimes A_6 \otimes A_5 \\ & A_5 \otimes A_5 \otimes A_6 \otimes A_6 \otimes A_5 \otimes A_6 \otimes A_6 \otimes A_6 \otimes A_6 \\ & A_5 \otimes A_5 \otimes A_6 \otimes A$$

We now know that we can multiply A_1 to A_5 in as few as 1344 multiplication operations!

But where do we put our brackets?

Matrix Chain Multiplication

$$\begin{split} M[i,j] = & \begin{bmatrix} 0 & \text{if } i = j \\ min_{i \leq k < j} & \{M[i,j] = M[i,k] + M[k+1,j] + p_{i+1}p_kp_j\} \\ A_1 \otimes A_2 \otimes A_3 \otimes A_4 \otimes A_5 \\ & \text{4x0 } 10x3 & 3x12 12x20 20x7 \\ p_0, p_1, p_1, p_2, p_3, p_4, p_4, p_5, p_6 \end{bmatrix} \end{split}$$

We now know that we can multiply A₁ to A₅ in as few as 1344 multiplication operations!

But where do we put our brackets?

We must focus on the selected k values

$$M[i,j] = \begin{cases} 0 & \text{if } i=j \\ \min_{1 \le k < j} \{M[i,j] = M[i,k] + M[k+1,j] + p_{i-1}p_kp_j\} \end{cases}$$

$$A_1 \otimes A_2 \quad A_3 \otimes A_4 \otimes A_5$$

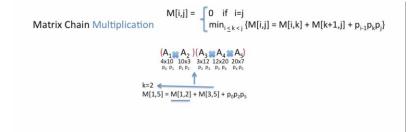
$$A_1 \otimes A_2 \quad A_3 \otimes A_4 \otimes A_5$$

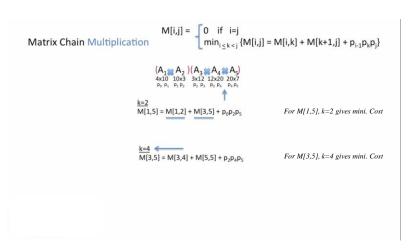
$$A_1 \otimes A_2 \quad A_3 \otimes A_4 \otimes A_5$$

$$A_1 \otimes A_2 \quad A_3 \otimes A_4 \otimes A_5$$

$$A_1 \otimes A_5 \quad A_5 \otimes A_6 \otimes A_6$$

$$A_1 \otimes A_5 \quad A_6 \otimes A_7 \otimes A_7$$





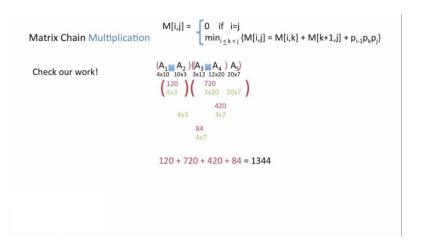
$$M[i,j] = \begin{cases} 0 & \text{if } i=j \\ \min_{1 \leq k < j} \{M[i,j] = M[i,k] + M[k+1,j] + p_{i-1}p_kp_j\} \end{cases}$$

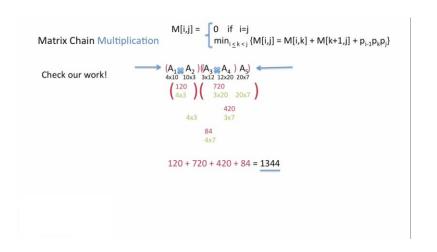
$$(A_1 \otimes A_2) \\ (A_1 \otimes A_3) (A_3 \otimes A_4) \\ (A_1 \otimes A_2) (A_3 \otimes A_4) \\ (A_1 \otimes A_3) (A_3 \otimes A_4) \\ (A_1 \otimes A_2) (A_3 \otimes A_4) \\ (A_1 \otimes A_3) (A_3 \otimes A_4) \\ (A_1 \otimes A_2) (A_3 \otimes A_4) \\ (A_1 \otimes A_3) (A_3 \otimes A_4) \\ (A_1 \otimes A_2) (A_3 \otimes A_4) \\ (A_3 \otimes A_3) (A_3 \otimes A_4) \\ (A_3 \otimes A_4) (A_4 \otimes A_4) \\ (A_4 \otimes A_4) (A_$$

$$\text{Matrix Chain Multiplication} \qquad \begin{array}{l} M[i,j] = \begin{cases} 0 & \text{if } i = j \\ \min_{1 \leq k < j} \{M[i,j] = M[i,k] + M[k+1,j] + p_{i-1}p_kp_j\} \end{cases} \\ \\ \text{Check our work!} \qquad \begin{array}{l} (A_1 \otimes A_2)([A_3 \otimes A_4]) A_5 \\ 4 \times 10^{-10.33} & 3 \times 12^{-10.20} & 20 \times 7 \\ 4 \times 3 & 10^{-10.20} & 3 \times 20^{-10.20} & 20 \times 7 \end{array}$$

Matrix Chain Multiplication
$$\begin{aligned} M[i,j] &= \begin{cases} 0 & \text{if } i=j \\ \min_{i \leq k < j} \{M[i,j] = M[i,k] + M[k+1,j] + p_{i+1}p_kp_j \} \end{aligned}$$
 Check our work!
$$\begin{aligned} (A_1 \otimes A_2) &((A_3 \otimes A_4) & A_5) \\ 4 \times 10 & 10 \times 3 & 3 \times 12 & 12 \times 20 & 20 \times 7 \\ 4 \times 3 & 0 & 3 \times 7 \end{aligned}$$

$$\begin{aligned} (A_2 \otimes A_4) &(A_3 \otimes A_4) &(A_5 \otimes A_4)$$





Thanks

Matrix Chain Multiplication Algorithm

Complexity On?

```
Matrix-Chain-Order(p)
                                                                     \label{eq:p0p1p0p1p0p1} \begin{array}{ll} \mbox{\ensuremath{/\!/}} p_0, p_1, p_2, ... = \mbox{Sizes of Matrices} \\ \mbox{\ensuremath{/\!/}} n & = & \mbox{No. of Matrices} \end{array}
            n = length[p] - 1
            for i = 1 to n do
                         m[i,i] = 0
            end-for
3
            for len = 2 to n do
                                                                       // len is the chain length.
4
                         for i = 1 to n - len + 1 do
5
                                      j = i + len - 1
6
                                     m[i, j] = Infinity
                                                                                    // Determine k for Optimal Soln
                                     for k = i to j - 1 do
                                                  q = m[i,\,k] + m[k+1,\,j] + p_{i\text{-}1}p_kp_j
                                                  if q < m[i, j] then
                                                               m[i, j] = q
11
                                                               s[i, j] = k
                                                   end-if
                                                                                     // s[i,j] saves the splitting pt. i.e. k
                                     end-for
            end-for end-for
12
            Return m and s
```