Matrix-chain Multiplication

Manuel Serna-Aguilera

Problem

Given sequence of matrices

$$A_1 \times A_2 \times A_3 \times A_4,$$

with dimensions

$$A_{5\times 4} \times A_{4\times 6} \times A_{6\times 2} \times A_{2\times 7},$$

we want to minimize the cost of multiplying the matrices.

Finding Cost

Given two matrices A and B, to find the cost or number of scalar multiplications of the matrix product $A \times B = C$, multiply the dimension of the rows and columns of A with the dimension of the columns of B. For example, given

$$A_{5\times 4} \times B_{4\times 3} = C_{5\times 3}$$

The number of scalar multiplications would be: $5 \cdot 4 \cdot 3 = 60$. Using a bottom-up approach, we start with single matrices, then move up to multiply pairs, and then sets of three, and so on, to get the cost of the entire sequence.

Now, create tables m and s with row i and column j,

- m: at $m[i, j]$, the minimum cost is stored when multiplying matrices from A_i to A_j. We want to get to $m[1, 4]$, since we have 4 matrices.
- s: at $s[i, j]$, row i will give the sequence of the optimal ordering; basically, it tells us how to parenthesize the sequence of matrices from i to j.

m

	1	2	3	4
1				
2				
3				
4				

s

	1	2	3	4
1				
2				
3				
4				

For the procedure, we only need to fill out half of the table along the diagonal going top-left to bottom-right, as cases below are symmetrical.

Start at $m[i, j]$, where $i = j$.

- The cost is zero, just having one matrix by itself would not cost anything to multiply. This is trivial, but for this diagonal, fill in all zeros.

m

	1	2	3	4
1	0			
2		0		
3			0	
4				0

Move up to the next diagonal.

- $m[1, 2]$: multiply matrices A_1 and A_2.
  $(A_{5\times4}) \times A_{4\times6}$ gives a cost of $5 \cdot 4 \cdot 6 = 120$.
  Thus, $m[1, 2] = 120$.
  Place value of 1 in $s[1, 2]$ since we split the sequence at A_1.

- $m[2, 3]$: multiply matrices A_2 and A_3.
  $(A_{4\times6}) \times A_{6\times2}$ gives a cost of $4 \cdot 6 \cdot 2 = 48$.
  Thus, $m[2, 3] = 48$.
  Place value of 2 in $s[2, 3]$ since we split the sequence at A_2.

- $m[3, 4]$: multiply matrices A_3 and A_4.
  $(A_{6\times2}) \times A_{2\times7}$ gives a cost of $6 \cdot 2 \cdot 7 = 84$.
  Thus, $m[3, 4] = 84$.
  Place value of 3 in $s[3, 4]$ since we split the sequence at A_3.

This gives the updated tables m and s.

m

	1	2	3	4
1	0	120		
2		0	48	
3			0	84
4				0

s

	1	2	3	4
1		1		
2			2	
3				3
4				

Now we have calculated the cost of multiplying all pairs possible. Now lets calculate the cost of multiplying sequences with three matrices.

- $m[1,3]$: Consider two possibilities.

  1. $A_{5\times 4} \times (A_{4\times 6} \times A_{6\times 2}) = A_{5\times 4} \times A_{4\times 2}$
     Now, get the cost of A_1, a single matrix, plus the cost of $A_2 \times A_3$ (the inner multiplication), and getting the overall cost with this specific sequence. Recall $m[1,1]$ and $m[2,3]$ have already been calculated.

     $$m[1,1] + m[2,3] + 5 \cdot 4 \cdot 2 = 0 + 48 + 40 = 88$$

  2. $(A_{5\times 4} \times A_{4\times 6}) \times A_{6\times 2} = A_{5\times 6} \times A_{6\times 2}$

     $$m[1,2] + m[3,3] + 5 \cdot 6 \cdot 2 = 120 + 0 + 60 = 180$$

  Now, take the smaller cost, which is 88, thus $m[1,3] = 88$. Also, $s[1,3] = 1$ since A_1 was our cut-off.

- $m[2,4]$: Consider two possibilities again.

  1. $A_{4\times 6} \times (A_{6\times 2} \times A_{2\times 7}) = A_{4\times 6} \times A_{6\times 7}$

     $$m[2,2] + m[3,4] + 4 \cdot 6 \cdot 7 = 0 + 84 + 168 = 252$$

  2. $(A_{4\times 6} \times A_{6\times 2}) \times A_{2\times 7} = A_{4\times 2} \times A_{2\times 7}$

     $$m[2,3] + m[4,4] + 4 \cdot 2 \cdot 7 = 48 + 0 + 56 = 104$$

  Now, take the smaller cost, which is 104, thus $m[2,4] = 104$. Also, $s[2,4] = 3$ since A_3 separates $A_2 \times A_3$ from A_4.

That gives us the updated tables.

m		1	2	3	4
	1	0	120	88	
	2		0	48	104
	3			0	84
	4				0

s		1	2	3	4
	1		1	1	
	2			2	3
	3				3
	4				

Now, calculate the cost of the entire sequence, $m[1,4]$.

$$\begin{aligned}
m[1,4] =& \min\{m[1,1] + m[2,4] + 5 \cdot 4 \cdot 7, \\
& m[1,2] + m[3,4] + 5 \cdot 6 \cdot 7, \\
& m[1,3] + m[4,4] + 5 \cdot 2 \cdot 7 \\
=& \min\{0 + 104 + 140, \\
& 120 + 84 + 210, \\
& 88 + 0 + 70\} \\
=& \min\{244, 414, 158\} \\
m[1,4] =& 158
\end{aligned}$$

Since our solution comes from splitting $A_1 \cdot A_2 \cdot A_3$ from A_4, $s[1,4] = 3$. This gives our final tables m and s.

m		1	2	3	4
	1	0	120	88	158
	2		0	48	104
	3			0	84
	4				0

s		1	2	3	4
	1		1	1	3
	2			2	3
	3				3
	4				

Solution

This is the general formula for finding the ideal sequence to multiply matrices i to j, with dimensions for rows and columns labeled with d.

$$m[i,j] = \min_{i \le k \le j} \{m[i,k] + m[k+1,j] + d_{i-1} \cdot d_k \cdot d_j\}$$

Given n matrices, we will have $n+1$ dimensions. Here is the optimal parenthesization.

$$(A_1 \cdot A_2 \cdot A_3) \cdot (A_4)$$

Cut off the sequence at third matrix since $s[1,4] = 3$.

$$\big((A_1) \cdot (A_2 \cdot A_3)\big) \cdot (A_4)$$

Cut off the sequence at first matrix since $s[1,3] = 1$.

Thus, this is the sequence of matrices that requires the least number of scalar multiplications.

$$\big((A_{5\times 4}) \times (A_{4\times 6} \times A_{6\times 2})\big) \times (A_{2\times 7})$$