Matrix Chain Multiplication

Gene Yang

12/9/2024

Introduction

- At this point, we've probably all multiplied matrices before.
- Say we have a ton of matrices to multiply:

$$M = A_1 A_2 A_3 \ldots A_N$$

- Matrix multiplication is associative, so (AB)C = A(BC).
- How should the parentheses be placed in the right-hand side to get the minimum number of operations?

- Note: only considering multiplications since addition takes much less time in comparison (for a computer at least)

Example

$$\left(\begin{bmatrix}6 & 9 & 2\\7 & 1 & 8\end{bmatrix}\begin{bmatrix}5\\4\\3\end{bmatrix}\right)\begin{bmatrix}1 & 2\end{bmatrix}$$

Multiplying matrix #1 & #2

- $2 \times 3 = 6$ multiplications

Multiplying that with matrix #3

- $2 \times 2 = 4$ multiplications

Total 10 multiplications

$$\begin{bmatrix}6 & 9 & 2\\7 & 1 & 8\end{bmatrix}\left(\begin{bmatrix}5\\4\\3\end{bmatrix}\begin{bmatrix}1 & 2\end{bmatrix}\right)$$

Multiplying matrix #2 & #3

- $3 \times 2 = 6$ multiplications

Multiplying that with matrix #3

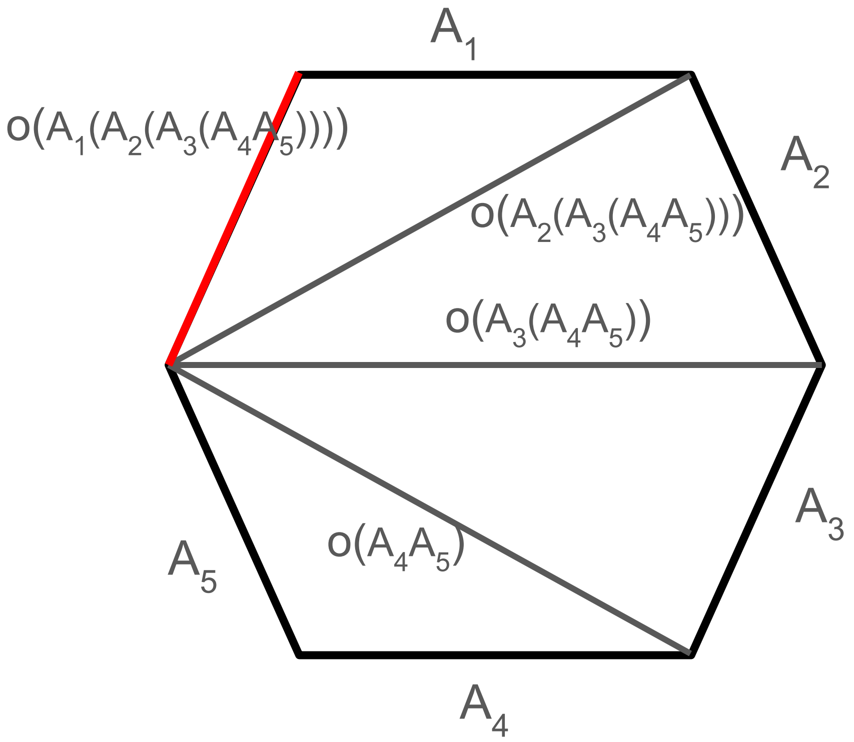
- $2 \times 3 \times 2 = 12$ multiplications

Total 18 multiplications

In general, if A is a $p \times q$ matrix and B is a $q \times r$ matrix, AB takes pqr operations to calculate.

Polygons

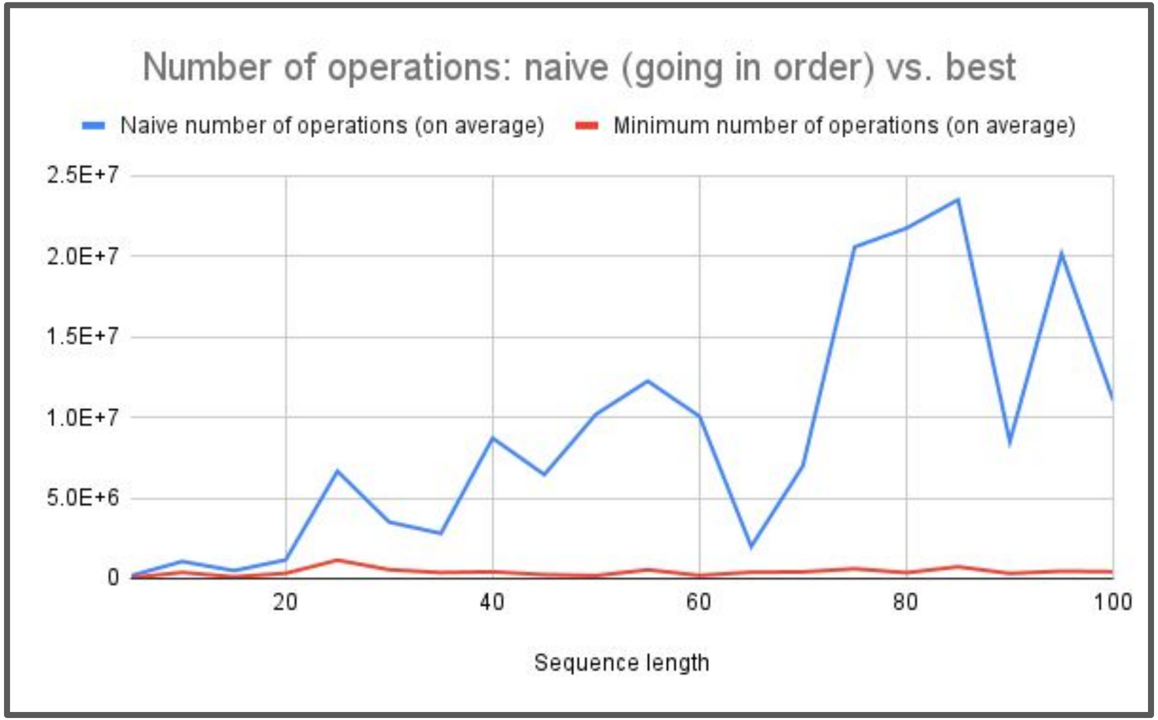
- A nice visualization with polygon triangulation:



1. Draw a (N + 1)-sided polygon. Place matrices on each edge in clockwise order.
2. Split the polygon into triangles, representing a pairing.
3. Each new edge formed by the triangle represents the # of operations for that pairing.
4. Total # of operations on (N+1)th side.

Some Tests

- A basic simulation of matrix chain multiplication I ran over the weekend.



Each point on the graph was found by averaging over 1000 runs.

Matrix dimensions could be anywhere from 1×1 to 100×100.

(bad code)

https://github.com/gyang0/Matrix-Chain-Multiplication

Applications

- Practically every technology using computers relies on matrices and multiplying them.
    - Computer graphics: lots of transformation & rotation matrices are multiplied to form scenes. Particularly true for 3D animations.
    - Image processing: each pixel can represent an element of a matrix, for example.
    - Machine learning: I don’t know much about this, but I think the model is based on large matrices.

- In real life, the matrices could be much larger. The size of the sequence would be larger too.

Links

Sources:

- https://ieeexplore.ieee.org/document/6553999/similar#similar
- https://ics.uci.edu/~eppstein/260/011023/

Stuff that might be interesting:

- O(N log N) algorithm by Hu & Shing: https://pdfs.semanticscholar.org/e8b8/40a921f7967b30ac161b3dd9654b27998ddb.pdf
- Catalan numbers: https://en.wikipedia.org/wiki/Catalan_number