Linear Algebra

ETH Zürich, HS 2024, 401-0131-00L

The Computer Science Lens

Multiplying Matrices and Solving Systems of Linear Equations Faster

Bernd Gärtner

October 17, 2024

The Cost of Gauss Elimination

Solving a system $A\mathbf{x} = \mathbf{b}$ of m equations in m variables using Gauss Elimination takes $O(m^3)$ operations.

Can we do it faster?

Yes, by doing something else faster: *matrix multiplication*!

Cost of Matrix Multiplication (square case $m \times m$)

$$AB = \underbrace{\begin{bmatrix} — & \mathbf{u}_1 & — \\ — & \mathbf{u}_2 & — \\ & \vdots & \\ — & \mathbf{u}_m & — \end{bmatrix}}_{A\text{, row notation}} \underbrace{\begin{bmatrix} | & | & & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_m \\ | & | & & | \end{bmatrix}}_{B\text{, column notation}} = \underbrace{\begin{bmatrix} \mathbf{u}_1 \cdot \mathbf{v}_1 & \mathbf{u}_1 \cdot \mathbf{v}_2 & \cdots & \mathbf{u}_1 \cdot \mathbf{v}_m \\ \mathbf{u}_2 \cdot \mathbf{v}_1 & \mathbf{u}_2 \cdot \mathbf{v}_2 & \cdots & \mathbf{u}_2 \cdot \mathbf{v}_m \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{u}_m \cdot \mathbf{v}_1 & \mathbf{u}_m \cdot \mathbf{v}_2 & \cdots & \mathbf{u}_m \cdot \mathbf{v}_m \end{bmatrix}}_{m^2\text{ scalar products}}$$

Scalar product of $\mathbf{u}, \mathbf{v} \in \mathbb{R}^m$: $\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \cdots + u_m v_m.$

Cost of computing AB:

- m multiplications per scalar product — m^3 multiplications in total
- $m - 1$ additions per scalar product — $m^3 - m^2$ additions in total

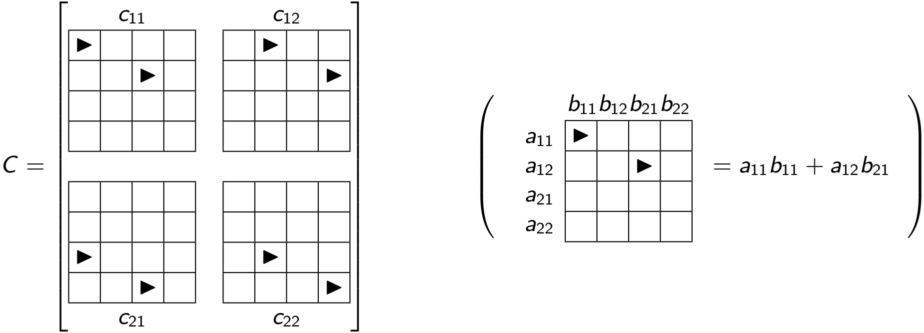
Can we do it with less operations (multiplications / additions)?

Volker Strassen (1969): yes! [Str69]

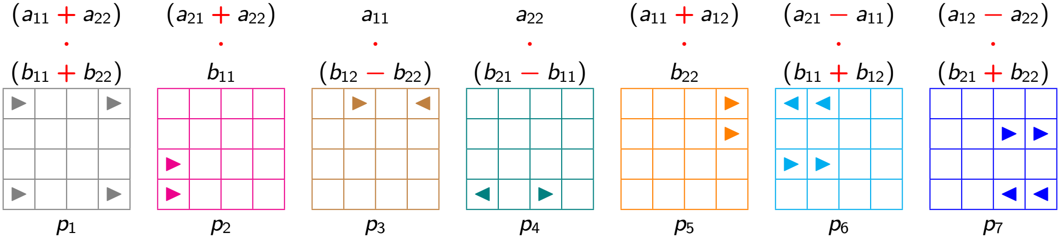
2×2, standard way: 8 multiplications $(\cdot)$, 4 additions $(+)$

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}}_{B} = \underbrace{\begin{bmatrix} a_{11} \cdot b_{11} + a_{12} \cdot b_{21} & a_{11} \cdot b_{12} + a_{12} \cdot b_{22} \\ a_{21} \cdot b_{11} + a_{22} \cdot b_{21} & a_{21} \cdot b_{12} + a_{22} \cdot b_{22} \end{bmatrix}}_{AB} = \underbrace{\begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}}_{C}$$

Graphically:



2×2, Strassen way: 7 multiplications $(\cdot)$, 18 additions $(+, -)$



$(a_{11} + a_{22}) \cdot (b_{11} + b_{22})$ — p_1

$(a_{21} + a_{22}) \cdot b_{11}$ — p_2

$a_{11} \cdot (b_{12} - b_{22})$ — p_3

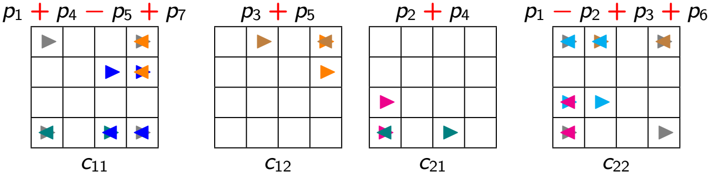
$a_{22} \cdot (b_{21} - b_{11})$ — p_4

$(a_{11} + a_{12}) \cdot b_{22}$ — p_5

$(a_{21} - a_{11}) \cdot (b_{11} + b_{12})$ — p_6

$(a_{12} - a_{22}) \cdot (b_{21} + b_{22})$ — p_7

Auxiliary terms: 7 multiplications, 10 additions (◀: product has negative sign)



$p_1 + p_4 - p_5 + p_7$ — c_{11}

$p_3 + p_5$ — c_{12}

$p_2 + p_4$ — c_{21}

$p_1 - p_2 + p_3 + p_6$ — c_{22}

Entries of $C = AB$: 8 additions

4×4: $\quad = 2 \times 2$ with *matrices* (2×2), not numbers!

$$\underbrace{\begin{array}{c} A_{11} \qquad\quad A_{12} \\ \left[\begin{array}{cc|cc} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ \hline a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{array}\right] \\ A_{21} \qquad\quad A_{22} \end{array}}_{A} \underbrace{\begin{array}{c} B_{11} \qquad\quad B_{12} \\ \left[\begin{array}{cc|cc} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ \hline b_{31} & b_{32} & b_{33} & b_{34} \\ b_{41} & b_{42} & b_{43} & b_{44} \end{array}\right] \\ B_{21} \qquad\quad B_{22} \end{array}}_{B} = \underbrace{\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}}_{\substack{2 \times 2 \text{ block} \\ \text{form of } A}} \underbrace{\begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}}_{\substack{2 \times 2 \text{ block} \\ \text{form of } B}}$$

$$= \underbrace{\begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{bmatrix}}_{\substack{2 \times 2 \text{ block} \\ \text{form of } AB}} \quad \leftarrow \text{check this!}$$

Standard way: 8 multiplications and 4 additions . . . of 2×2 matrices

Strassen way: 7 multiplications and 18 additions . . . of 2×2 matrices

Cost (how many operations on numbers?)

4×4:

Standard	Strassen
4^3 $\quad$ $4^3 - 4^2$	7 multiplications and 18 additions of 2×2 matrices

$$\begin{array}{|c|c|}\hline \multicolumn{2}{|c|}{(4\times 4)\cdot(4\times 4)}\\\hline \cdot & + \\\hline 64 & 48\\\hline\end{array} \qquad 7 \cdot \begin{array}{|c|c|}\hline \multicolumn{2}{|c|}{(2\times 2)\cdot(2\times 2)}\\\hline \cdot & + \\\hline 7 & 18\\\hline\end{array} + 18 \cdot \begin{array}{|c|}\hline (2\times 2)+(2\times 2)\\\hline + \\\hline 4\\\hline\end{array} = \begin{array}{|c|c|}\hline \multicolumn{2}{|c|}{(4\times 4)\cdot(4\times 4)}\\\hline \cdot & + \\\hline 49 & 198\\\hline\end{array}$$

8×8 (use 2×2 block form with 4×4 matrices as blocks):

Standard	Strassen
8^3 $\quad$ $8^3 - 8^2$	7 multiplications and 18 additions of 4×4 matrices

$$\begin{array}{|c|c|}\hline \multicolumn{2}{|c|}{(8\times 8)\cdot(8\times 8)}\\\hline \cdot & + \\\hline 512 & 448\\\hline\end{array} \qquad 7 \cdot \begin{array}{|c|c|}\hline \multicolumn{2}{|c|}{(4\times 4)\cdot(4\times 4)}\\\hline \cdot & + \\\hline 49 & 198\\\hline\end{array} + 18 \cdot \begin{array}{|c|}\hline (4\times 4)+(4\times 4)\\\hline + \\\hline 16\\\hline\end{array} = \begin{array}{|c|c|}\hline \multicolumn{2}{|c|}{(8\times 8)\cdot(8\times 8)}\\\hline \cdot & + \\\hline 343 & 1674\\\hline\end{array}$$

Cost (how many operations on numbers?)

$2^k \times 2^k$ (general formula; officially needs proof by induction):

Standard		Strassen	
$(2^k \times 2^k) \cdot (2^k \times 2^k)$		$(2^k \times 2^k) \cdot (2^k \times 2^k)$	
$\cdot$	$+$	$\cdot$	$+$
8^k	$8^k - 4^k$	7^k	$6(7^k - 4^k)$

Checking...

	Standard		Strassen	
	$\cdot$	$+$	$\cdot$	$+$
k	8^k	$8^k - 4^k$	7^k	$6(7^k - 4^k)$
1	8	4	7	18
2	64	48	49	198
3	512	448	343	1674

Is Strassen ever faster?

Let's count total number of operations!

	Standard	Strassen	Strassen − Standard
	$\cdot/+$	$\cdot/+$	$\cdot/+$
k	$2 \cdot 8^k - 4^k$	$7 \cdot 7^k - 6 \cdot 4^k$	$7 \cdot 7^k - 2 \cdot 8^k - 5 \cdot 4^k$
1	12	25	13
2	112	247	135
3	960	2017	1057
$\vdots$			
9	268173312	280902385	12729073
10	2146435072	1971035287	−175399785
11	17175674880	13816121377	−3359553503
$\vdots$			

For 1024 × 1024 (and larger) matrices, Strassen is faster!

$m \times m$

There is a version of Strassen's algorithm that works for all m (not only $m = 2^k$).

	Standard	Strassen
	$\cdot / +$	$\cdot / +$
m	$\approx 2m^3$	$\approx 4.7m^{2.8}$

$\log_2(7) = 2.80735\ldots$

Strassen's algorithm started an ongoing race for better exponents than 2.8.

The "2.37 . . . family": since 1990

- $\approx cm^{2.376}$ (c some large constant) [CW90]
- Latest: $\approx c'm^{2.371866}$ (c' some even larger constant) [DWZ23].

Some people believe that $\approx Cm^2$ (C a super large constant) is possible, but nobody knows how to get there.

Applications

How important is matrix multiplication? Very important!

Many matrix problems (some of them introduced later in this course) can be reduced to matrix multiplication and therefore be solved faster using the algorithms by Strassen and his successors.

Important for us:

- Computation of the inverse with $\approx cm^{2.37\cdots}$ operations
- Proof by reduction of inverse computation to matrix multiplication [THCS22]
- Since solving $A\mathbf{x} = \mathbf{b}$ reduces to inverse computation ($\mathbf{x} = A^{-1}\mathbf{b}$), we can also solve systems of linear equations faster than using Gauss elimination.
- Also true (but more difficult) if system has no unique solution / A^{-1} does not exist [Sch72, KG85]
- Big theoretical improvement over $\approx cm^3$ that we get from Gauss elimination
- Practical improvements only for impractically large m

References

Don Coppersmith and Shmuel Winograd.
Matrix multiplication via arithmetic progressions.
Journal of Symbolic Computation, 9(3):251–280, 1990.
https://doi.org/10.1016/S0747-7171(08)80013-2.

Ran Duan, Hongxun Wu, and Renfei Zhou.
Faster matrix multiplication via asymmetric hashing, 2023.
https://arxiv.org/abs/2210.10173.

Walter Keller-Gehrig.
Fast algorithms for the characteristics polynomial.
Theoretical Computer Science, 36:309–317, 1985.
https://doi.org/10.1016/0304-3975(85)90049-0.

A. Schönhage.
Unitäre Transformationen großer Matrizen.
Numerische Mathematik, 20(5):409–417, 1972.
https://doi.org/10.1007/BF01402563.

Volker Strassen.
Gaussian elimination is not optimal.
Numerische Mathematik, 13(4):354–356, 1969.
https://doi.org/10.1007/BF02165411.

Ronald L. Rivest Thomas H. Cormen, Charles E. Leiserson and Clifford Stein.
Introduction to Algorithms, Fourth Edition.
The MIT Press, 2022.
https://mitpress.mit.edu/9780262046305/introduction-to-algorithms/.