What is...machine learning in mathematics - part 2?

Or: AI and matrix multiplication

Standard matrix multiplication is not optimal!

$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \times$	$\begin{pmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} \end{pmatrix} = \begin{pmatrix} \mathbf{A} \\ \mathbf{C} \end{bmatrix}$	E+BG AF+BH E+DG CF+DH
Trick:		
$P_1 = A \cdot (F - H)$	$P_{\!\scriptscriptstyle 5} = (A \!+\! D) \!\cdot\! (E \!+\! H)$	$AE + BG = P_5 + P_4 - P_2 + P_6$
$P_2 = (A+B) \cdot H$	$P_{\!6}=(B\text{-}D){\cdot}(G\text{+}H)$	$AF + BH = P_1 + P_2$
$P_3 = (C+D) \cdot E$	$P_7 = (A-C) \cdot (E+F)$	$CE+DG = P_3 + P_4$
$P_4 = D \cdot (G - E)$		$CF+DH = P_5 + P_1 - P_3 - P_7$

► Standard 2-by-2 matrix multiplication needs eight multiplications

Strassen We only need seven

Observation The seven multiplications are not trivial to find

Divide and conquer



Recursion Safe one operation each step while multiplying 2ⁿ-by-2ⁿ matrices

Standard multiplication Needs n^3 operations (3 = $\log_2 8$)

Strassen multiplication Needs $n^{2.81}$ operations $(2.81 \approx \log_2 7)$

Why not k-by-k matrices?



- Idea Why don't we use large matrices for the recursion (difficult remember how weird Strassen's formulas look!)
- Example If for 3-by-3 matrices we safe 7 operations, then we are down to $2.73 \approx \log_3 20$
- Idea Use reinforcement learning, let AI play a game and reward saving operations

Deepmind's AlphaTensor found a 5-by-5 algorithm with 76 operations

 $\log_5 76 \approx 2.69$

The neutral network used is not a traditional one but used reinforcement learning

- ► They actually discovered several other algorithms as well
- ► Here is the list Don't look ☺

no na na na na ha	
and and and and and the second	or has had a second reason of the
NO NO NO NO NO NO	o ha ha 🖂 👘 na na na na na
NO FO NO NO FO	a ha ha ina na na na na na
And the Price Republic	
And the second distance for the second	
And second and a first state of	
And the second address of the	
A REPORT OF A DESCRIPTION OF A DESCRIPTI	
And the mental Rent K. K.	
Destination and No. 1944	
And the second distance of	
And second and the sheet	
A contract of the second s	
Designed and a second state of the second	
Sector Sector	
And the second dealers'	
And the second dealers of	
An address of the second s	
An el concerno des	
The relative trace that the trace that the trace that the	
Autor Buthathat	
And one Read Academic	
The effect of the second data where the first	
The manufacture of the second s	
An electric constraints and the electric constraints are the electric	
Autor and Marthathat	
Autor March 1994	
An Arrest Revenue	
the street and the data	
And the second stand date that the standard	
Second Phene Render	
To the Ballballand	
The Title and a stand built	
and the second se	
the second second second second second	
Contraction of the second s	
Inclusion of the later	
And an	
The fillence is a construction of the state of the state of the	
Tester stored block of a back a book a book of	
have been a	
Automatica Automatica Automatica	
The street many start that has the the the street had	
Test methods and	
The first sector is a first starting the short of	
Dealers and Barbards also the thirth of	
Designs encourse and an encourse design back and	
The second date is the she she she she	

Actually...



- ► There are actually faster algorithms
- ► Most of the newer ones use different methods
- Problem Most of them are not of practical importance

Thank you for your attention!

I hope that was of some help.