Mathematics in the modern world

E. Zelmanov

Zhejiang Normal University

On Mathematical Aspects of AI

Never, througout history, has Mathematics had such a profound impact on our lives as it does today. I should say: welcome to the bright new world of Mathematics. We live at the time of the Informational Revolution, which has followed others that came before, such as the Industrial Revolution and the Agricultural Revolution; all, by the way,with a high impact on the workforce, for better or for worse. And this revolution has been brought about by Mathematics.

On one hand Mathematics is driven by inner logic. Such problems as the Riemann Conjecture and a few other problems have been at the center of mathematics 100 years ago and they are still at the center.

On Mathematical Aspects of AI

On the other hand, modern technology has formulated mathematical problems of unprecedented scale and importance. I will formulate only 3 problems that, in my opinion will stay at the center of mathematical applications at least till the end of the century.

The first problem is : mathematical understanding of the Deep Learning Algorithm.

I'll tell you a story that happened a few years ago in California. There was a very serious movement there to eliminate or water down mathematics in school. It is a difficult subject, many students get bad grades, which then prevents them from going to good universities.

On Mathematical Aspects of AI

The issue was much discussed and then (!) : all the CEOs of the big AI companies wrote an open letter stating that artificial intelligence grew out of mathematics. AI speaks the language of vectors and matrices$\cdots$ the language of Mathematics. In short, no Mathematics - no AI.

Here is the letter. Please, read it. I won't be able to say it better.

Open Letter

Artificial Intelligence is poised to transform society as we know it. To be prepared for this future, it is imperative we educate our future workforce with the knowledge to build and deploy AI technology. Core mathematical concepts from algebra, calculus, and probability lie at the heart of modern AI innovation. As such, engaging in the development of these technologies requires that students start with strong mathematical foundations.

Open Letter

While today's advances might suggest classic mathematical topics like calculus or algebra are outdated, nothing could be further from the truth. In reality, modern AI systems are rooted in mathematics, making a strong command over math necessary for careers in this field. The algorithmic backbone of deep learning—gradient descent—exemplifies this connection by combining calculus with (linear) algebra.

Vectors and matrices are the building blocks of neural networks, and modeling growth on a logarithmic scale is fundamental to the science of neural network training. Trigonometric and Pythagorean identities, rather than being "outdated," underlie crucial tools from data science, including the Fourier transform and least squares algorithms. Studying these core topics in high school is the best preparation for later specialization in machine learning, data science, or any STEM field, and generally we would much rather hire students who have mastered fundamentals than those who have a shallow familiarity with the latest tools or software.

Open Letter signature

- Sam Altman. CEO, OpenAI
- Shyamal Anadkat. Applied AI, OpenAI
- Samy Bengio. Senior Director of Machine Learning Research, Apple
- Sebastien Bubeck. VP GenAI Research, Microsoft
- Bryan Catanzaro. VP Applied Deep Learning Research, NVIDIA
- Adam Cheyer. Co-founder, Siri; VP of AI Experience, Airbnb
- Bill Dally. Chief Scientist and Senior Vice President of Research, NVIDIA
- Jeff Dean. Chief Scientist, Google DeepMind and Google Research; ACM Prize in Computing, IEEE John von Neumann Medal
- Jonathan Frankle. Chief Scientist, MosaicML; Chief Scientist Neural Networks, Databricks
- Ali Ghodsi. CEO and Co-founder, Databricks
- John Giannandrea. SVP Machine Learning and AI Strategy, Apple
- Boris Ginsburg. Senior Director, NVIDIA
- Steve Golik. CEO and Co-Founder, Juice Labs
- Ramin Hasani. CEO and Co-Founder, Liquid AI
- Urs Hölzle. Google Fellow and Former SVP of Technical Infrastructure, Google
- Jeff Huber. Former SVP, Google; Founder, GRAIL & Triatomic Capital (AI fund); Chairman, College of Computing, Data Science & Society (CDSS), UC Berkeley
- Michael Kagan. CTO, NVIDIA
- Branislav Kisačanin. Co-founder and Chief Scientist at Institute of AI R&D of Serbia; Sr. Architect at NVIDIA; Instructor at AwesomeMath

Open Letter signature

- Yann LeCun. Chief AI Scientist, Meta; Jacob T. Schwartz Professor of Computer Science & Data Science, NYU; Turing Award Laureate; NAS, NAE
- Edo Liberty. CEO and Founder, Pinecone
- Curtis Liu. CTO and Co-founder, Amplitude
- Ming-Yu Liu. VP GenAI Research, NVIDIA
- Peyman Milanfar. Distinguished Scientist, Google Research
- Vedant Misra. Research Engineer and Technical Lead — Gemini Core Pretraining, Google DeepMind
- Mira Murati. CTO, OpenAI
- Elon Musk. X
- S.Muthu Muthukrishnan. VP Sponsored Products, Amazon Advertising
- Mekka Okereke. General Manager, Google Play Apps
- John Overdeck. Co-chair and Co-founder, Two Sigma
- Jakub Pachocki. Chief Scientist, OpenAI
- Swami Sivasubramanian. VP Data and Machine Learning Services, Amazon Web Services
- Spenser Skates. CEO, Amplitude
- Ion Stoica. President and Co-founder, Anyscale; Co-founder, Conviva; Executive Chairman and Co-founder, Databricks
- Andrew Sutherland. Founder, Quizlet
- Richard Tang. Senior Engineer, Cisco Systems
- Avery Wang. Co-founder and Inventor, Shazam
- Greg Yang. Co-founder, xAI
- Matei Zaharia. CTO and Co-founder, Databricks
- Denny Zhou. Founder and Lead of the Large Language Model Reasoning Team, Google DeepMind

On Mathematical Aspects of AI

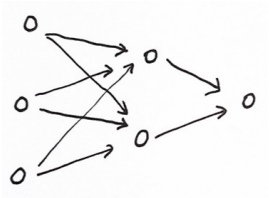
Direct Problem: Compute a function

$g(x) = y$.

Inverse Problem:We know that $g(x_i) = y_i, 1 <= i <= m$. Find a fuction $g(x)$.

Polynomial approximation.

Artificial deep neural network



Weight 2×3

Plus we choose a bias vector of length 2

$\rho(x) = \max(0, x)$ activation function

Training Process

$\{x^{(i)}, y^{(i)}\}_{i=1,2,\ldots,m}$ samples, $x^{(i)}$ have length n_0, $y^{(i)}$ are numbers.

Choose x, and n-tuple of length n_0.

$$x \to W_1 x + b_1 \to \rho(W_1 x + b_n).$$

Then again multiply on the left by W_2, add b_2 and apply ρ.

And so on.

After d steps we get a function $f(x)$.

We want: $f(x_i) = y_i$, $1 \leq i \leq m$.

But it depends on matrices $W_1, \ldots, W_d$ and bias vectors $b_1, \ldots, b_d$.

We minimize

$$\frac{1}{m} \sum_{i=1}^{m} \left(f(x^{(i)}) - y^{(i)} \right)^2$$

via gradient descent.

On Mathematical Aspects of AI

Everything is too large $\Longrightarrow$ stochastic gradient descent.

1. Linear Algebra
2. Optimization
3. Statistics.

Many Open Problems:

Why gradient methods work? The problem is no convex.

Nonpolynomial approximations?

Complexity?

Choice of samples?

"Unavoidable" Problem:

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = l_1 \\ \qquad\qquad \vdots \\ a_{n1}x_1 + \cdots + a_{nn}x_n = l_n \end{cases}$$

Gaussian elimination algorithm: complexity n^3

If n is huge then n^3 is not practical.

On Mathematical Aspects of AI

I should remark that unlike, say, The Riemann Conjecture this problem does not touch on the heart of Mathematics. But, Deep Learning is currently the most used mathematical algorithm. Any advance in understanding will bring big advances in efficiency.

The second problem is Cybersecurity.This story starts about 200 years ago.

Évariste Galois (1811-1832), a French teenager.

Quadratic equation:

$$x^2 + ax + b = 0$$

$$x_{1,2} = \frac{-a \pm \sqrt{a^2 - 4b}}{2}$$

Cubic equation: Cardano

Equation of degree 4: Ferrari

5? 6?

P. Ruffini (1799) : an incomplete proof that a formula for equations of degree 5 does not exist.

N. H. Abel (1824) : correct proof!

Symmetries in Mathematics

E. Galois (building on ideas of Lagrange): groups of symmetries explain everything.

20th Century: Elementary Particles.

Groups of symmetries of elementary particles.

" I dont know if God exists, but if he exists, then he knows Group Theory."

Mathematical Examples

E. Galois: finite fields.

$\mathbb{Z}$ integers, p a prime number

$$a = qp + r,\ 0 \leq r < p - 1$$

r is the remainder of a modulo p.

$F(p) = 0, 1, 2, \ldots, p - 1$

$i + j =$ remainder of $i + j$ modulo p

$ij =$ remainder of ij modulo p

$2 + 4 = 1\ mod\ 5,\ 2 \cdot 4 = 3$

All laws are the same as in real or complex numbers, we can even divide by nonzero remainders.

Mathematical Examples

As good as real numbers, but ...finite.

Message: a sequence of 0 or $1's$. Remainders modulo 2.

$V = F(2)^n$ vector space over F(2).

Code = a fixed subspace of $F(2)^n$.

$$W \subset V$$

We transmit only elements from W. If $a \in V \setminus W$ was received then there was a mistake.

Usually: find closest element to a in W.

Mathematical Examples

Distance: number of positions where two vectors are different (one has 0, another one has 1).

Hamming weight $H(W) =$ minimal distance between distinct elements in W.

Good Code: large Hamming weight.

Example:

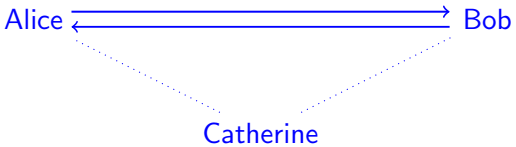
Binary Golay Code, $V = F(2)^{24}$, $\dim W = 12$, $H(W) = 8$.

Applications.

1. Voyager missions, pictures of Jupiter and Saturn;
2. USA standards for automatic link establishment in High Frequency radio specifies the use of the Golay Code;
3. In Mathematics: Aut(Golay Code) $= M_{24}$; The Monster Group; Conformal Field Theory in Mathematical Physics; Sphere Packing, etc.

Applications: Cryptography

Public Cryptography.



Diffie-Hellman, key exchange, 1976

M. Williamson, 1974, classfied.

Applications: Cryptography

RSA Rivest, Shamir, Adleman, 1978

$F(p) = \{1, 2, \ldots, p-1\}$

$F(p)^* = \{1, 2, \ldots, p-1\}$ is a cyclic group with respect to multiplication, g a generator of $F(p)^*$.

Alice		Bob
a secret		b secret
(a, g)	$\rightarrow$	g^a
g^b	$\leftarrow$	(b, g)
$(g^b)^a$	$=$	$(g^a)^b$

Alice and Bob both know g^{ab}. Catherine knows: g^a, g^b, g. To find g^{ab} she needs to know a, b.

Applications: Cryptography

The Problem of Discrete Logarithm.

AES (Advanced Encryption Standard, 2001).

Designed: J. Daemen, V. Rijmen

1998(Rejndael)

Bit $=$ 0 or 1

Bite $=$ 8 bits $=$ elements of $F(2)^8$

$|F(2)^8| = 2^8 = 256$

There exists a field $F(2^8)$ of order 256 (Galois).

Applications: Cryptography

We send: a 4 x 4 matrix over $F(2^8)$

$$(a_{ij})_{1<=i,\ j<=4}$$

Mix rows and columns, add keys + S-box operation:

$$(a_{ij}) \rightarrow (a_{ij}^{-1}), \textit{where}\, 0^{-1} = 0.$$

$F(2^8)^*$ is a cyclic group with generator a of order 255, $a^{-1} = a^{254}$.

Applications: Tomography

J. Radon, 1917, Functional Analysis.

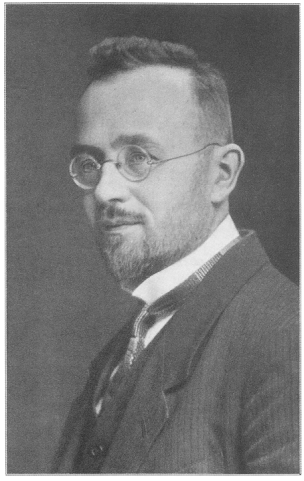
Recover a function from integrals over straight lines (the inverse problem to integration).

1970, Cormak, Haunsfield: TOMOGRAPHY.

However, too many X-rays are not good for patients.

Ultrasound Tomography is weaker.

Why? X-rays are strong. They go straight.



Johann Radon (1887-1956)

Applications: Tomography

Ultrasound is weaker, waves propagate along certain curves.

Recovering functions from integrals over curves is a much more difficult mathematical problem. There are no easy and direct for-mulas similar to Radon's formulas for straigh -t lines.

Turing Machines

M. Dehn (1906) $G = \langle a_1, \ldots, a_m | r_1 = 1, \ldots, r_s = 1 \rangle$

$v(a_1, \ldots, a_m) \stackrel{?}{=} w(a_1, \ldots, a_m)$.

Algorithmic Problems

Definition of an algorithm, Turing machines.

Mathematics has two sides

In experimental sciences the criterion of truth: repetition of the experiment.

In Mathematics: Her Majesty Proof.

John von Neumann: this concept changed several times during my lifetime. Still... it is amazingly stable. Euclid's proofs are still proofs.

What is a proof? A proof is what is considered to be a proof by all mathematicians;

Therefore they should better agree about it.

Purpose of proof

What is the purpose of a proof?

Understanding.

A proof can be beautiful or ugly.

What is beauty in Mathematics?

(i) a simple statement, a complicated and deep proof (like The Fermat Problem),
(ii) unexpected ideas, coming sometimes from a different area,
(iii) generality, when the same idea, sometimes in different forms, shows in different contexts.

A "golden standard of beauty in Mathematics": Galois Theory.

Beauty in Mathematics

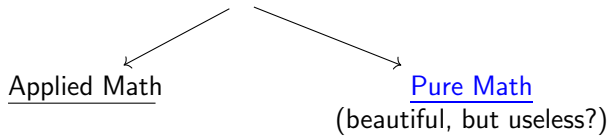
Students in Art Schools copy great paintings and learn what is great.

Tastes evolve in time: I am not sure that Shostakovich's music would be appreciated at the time of Mozart.

Mathematics is an elitist Art. Yet, it is the best supported Art. The biggest employer of specialists in Algebra & Number Theory is the National Security Agency of the USA.

Conclusions

Early 20th Century: Abstract Revolution

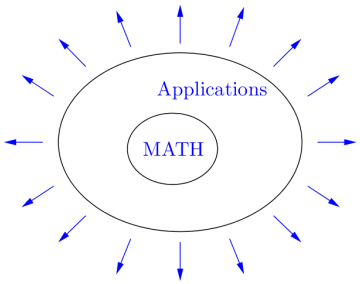


Can we support only Applied Math?

No, because Galois, Radon, Jim Simons would count as pure mathematicians.

Math is like a plant: all parts are related and feed each other. If you cut "unneeded parts" you kill the plant.

Conclusions



New Challenges: "Big Data", Machine Learning

The 21st Century

E. Wigner: unreasonable effectiveness of Mathematics.

Thank you!