The Unreasonable Expectation of Mathematics

August Lau

Abstract. Mathematics has been highly effective in its application of formalism to the real world. It spans from physical science to data science. Mathematical algorithm affects our everyday life. Mathematization (converting data to equations or mathematical forms) has been very successful. It has led to comments like "the unreasonable effectiveness of mathematics," but we might have "unreasonable expectation of mathematics."

Introduction. We will examine the expectation of mathematics in its relationship to the overwhelming amount of data. Since we compress data into equations or other forms, it loses its details. So mathematization works well in a "pure or noise-free" environment. Seismic imaging is an illustration of "noisy" data. It is important for us to interpret data.

We will confine our discussion using the theoretical background of topology/semigroup and the practical background of seismic imaging. We have collected big datasets of seismic data from the field in exploration of oil and gas. Hopefully, the two diverse backgrounds, theoretical vs practical, are sufficient to allow us to make some concluding remarks.

<u>"The Unreasonable Expectation of Mathematics"</u>

Data convinces us that we have "An unreasonable expectation of mathematics" and why:

Data convinces us that a solution to an inverse problem is not sufficient. In the real world, we are given data and we have to find the answer. Add as many terms as we like… yet it does not fit the data perfectly.

We celebrated the success of gravitational equation by dropping balls at the Tower of Pisa. If we drop a feather instead, mathematics will be powerless to predict the landing location of the feather. Mathematics governing gravity is an important success. But it is only applicable to "pure" situations: ball is bigger and heavier than air molecules. Once we have billions of interactions like air molecules, mathematics is truly challenged. If we go from physics to biology, we have many observations in biology; thus far, they defy mathematization.

Experiencing Mathematics

Unless we **live and breathe data**, we do not know the limitation of mathematics. Unless we **live and breathe mathematics**, we do not know the limitation of physical reality. We tend to think of mathematics as an austere subject devoid of emotion. When we prove a theorem, it is quite emotional and experiential. We experience **epiphany** even though the proof is cold and rigorous. By the same token, we experience **epiphany** when we can directly perceive data after processing with mathematical tools.

In order to make the data problem more tractable, data can be decomposed into a simple part and a complex part. The simple part can be described by models and mathematics. But the complex part is left unexplained by formal methods, but can be interpreted.

Academic vs industry problems:

In academic research, we can choose the problem that we want to solve. In industry, problems exist and we cannot simply ignore them. In exploration for oil and gas, we have an area to investigate to determine whether to drill a well or not. Seismic imaging helps us to make a decision, but we cannot wait for seismic problem to be completely solved before we drill. It is time critical in what we do in industry.

Typical problem:

You are given a bunch of data. Classical answer is to define the governing equation. We might include other operators like neuralnet (popularized by AI). Regardless, data has high dimension and has more detail. Data is compressed into a few variables or low dimensional space for practical implementation in computer algorithms. We have control over lab experiments and can describe them accurately. However, it is still a challenge to go from microscopic behavior to macroscopic behavior with real data.

Data convinces us that interpretation of residuals is important:

Residual is what the equation cannot explain. It is the difference between real data and modeled data. Residual indicates the magnitude of "failure" of the mathematical method. Yet, residual is useful for interpretation. Residual in seismic imaging can be interpreted and the residual is geologically meaningful. It is not just incoherent noise. In the past, we just ignore it and not bother to interpret it. However, successful seismic imaging is combining mathematics and interpretation.

Who is the end user?

Humans are the end users in major decision making. Humans are interpreters of data with aid from computer algorithms. Representation and processing of data are critical to interpretation.

Intuitive feelings about mathematics:

1. A mathematical algorithm will continue to be an anxious object.
2. Formalism proves rigorously that formalism itself is incomplete.
3. Data is more abundant than formal structures.
4. The goal of formalism is to understand data.

<u>Mathematics and three attributes:</u>

There are some concerns about the high expectation of mathematics. To illustrate, we will discuss three attributes. Can mathematics achieve these three attributes: accuracy, infinity, prediction?

Mathematics and accuracy:

Mathematics can explain the simple part of data, but not the complex part of data, which is the residual. Mathematics is not accurate in data fitting. It is difficult to see that mathematics is the ultimate explanation when we see residuals all the time. Even when we reduce it to one parameter like exp(x), the exponential keeps increasing. Mathematically, it has an inverse like exp(-x). In real data, like seismic data, exp(x) is not invertible.

Mathematics and infinity:

It is interesting how abstraction, like the mathematics of Cantor, almost let us feel as though we comprehend infinity. In fact, mathematically, there are more and more levels of infinity. It starts with a countable set of integers, and could increase to larger and larger uncountable infinite sets.

The flip side of Cantor's impact is the Cantor set, which challenges the notion that dimension is an integer. This has been popularized as fractal. This unusual notion of Hausdorff dimension is decimal rather than integer. We understand that space is 3-dimensional regardless of inductive or Hausdorff dimension. However, the inductive dimension of Cantor set is 0, but the Hausdorff dimension is between 0 and 1.

Mathematicians struggled long ago to accept abstract objects like infinite sets and non-integer dimensions. Most mathematicians now accept the reality of abstract objects. The digital world has pushed us further into non-physical reality. Subsequently, mathematics did expand and went beyond physical objects. Therefore, we can say that we somewhat understand infinity.

Mathematics and prediction:

Abstract mathematics has had a significant impact in the digital world. The digital world is real when we survey people who use network like social media or banking. Money is just a number in the bank without physical paper. Digital data is indispensable in our daily lives. The digital world is real in the sense that it affects emotions and physical well-being.

Connectivity (network) changes how we do mathematics. The prediction of weather in the future affects our current behavior. We decide to drive a car or to walk based on future weather prediction. However, our current decision to drive a car creates pollution which affects future weather. Consequently, the future prediction changes our current behavior. Now the future time prediction is connected to current time. What do we mean by prediction in the digital age when future prediction affects current decision? If billions of people watch weather prediction, how do billions of decisions affect the future?

Compressive mathematics and Generative mathematics:

Classical approach is to compress data into equations or low dimension space so that they are more manageable. Compression yields to lower computational effort. But is that really how data works? Living organisms are quite reproductive and they could grow at an exponential rate. It is generative and extensive in the sense that the size gets bigger and bigger.

Cantor certainly made a type of generative mathematics famous. Cantorian mathematics is generative (bigger and bigger infinity) and not reductive. Given a set S, we can generate the power set P(S) by collecting subsets of S. Taking the power set P(S), we can generate a new power set P(P(S)), etc. In the theoretical domain, it gets bigger and bigger. Another simple example of generative mathematics is semigroup of words which is just a string of alphabets. If we have 26 alphabets, we could generate finite arbitrary combination of alphabets, e.g., abc, bcz,wxhk, gubty,… which form an infinite semigroup by concatenation of words.

Conclusion:

>> Mathematics has great impact on physical science and data science
>> Digital reality could be non-physical, but data can motivate new mathematical research
>> Mathematicians should not be constrained by the comfort zone of reductionism (clean)
>> Mathematics is both rigorous and empirical (using both reductionism and empiricism)
>> Mathematics could be generative and infinite dimensional with no limit (unbounded)
>> Group theory has invertible operators, but semigroup has non-invertible operators
>> Non-invertible operators are difficult to deal with but realistically model real data
>> Mathematics favors a closed system, but it could still generate bigger and bigger sets

Mathematics is unreasonable and is beyond our expectation:

A mathematician has at least two roles: one role is doing the traditional mathematics of proving theorems and solving problems; another role is to invent new mathematics. An important invention was calculus, which went beyond data fitting by algebra. The motivation is from the **physical world**. Since then, mathematics has trended toward quantitative methods.

The motivation for new mathematics is driven by the **digital world** with data collected every second. Data mathematics could include qualitative methods like topology (homology or cohomology) or semigroup (non-invertible operators). Computers have made empirical investigation feasible. The new mathematics might be invented through collaboration with other disciplines. **It is scary to leave our mathematical home base and listen to strangers. But it could set us free.**

"The essence of mathematics lies in its freedom." Georg Cantor

Appendix:

It is difficult to predict where mathematics will go. But two mathematical theorems could shed some light on the future.

Weierstrass Approximation Theorem
A function can be approximated by **polynomials** (roughly speaking)

Using this theorem, we can define explicitly a polynomial relationship mapping of input directly to desired output. Polynomial is generalized to power series, rational functions, continued fractions, etc. It is not clear how to “clean” or “pre-condition” real data from the field to fit polynomials.

Universal Approximation Theorem
A function can be approximated by **neural network** (roughly speaking)

This is a fundamental theorem to machine learning. NN (neural network) is generalized to multi-layer NN. Instead of construction like polynomial, an “activation” function like sigmoidal function is a smooth form of an impulse function (on or off) since impulse function is non-differentiable. NN could be trained to learn “physical” equations like wave equation.

Changing view in data processing: Is data operator? Is operator data? Are data and operator the same? The changing view of data processing with the aid of theoretical mathematics has blurred the boundary of operator and data.

In linear algebra, a solution vector or a data vector is a matrix. Operator is a matrix. So data and operator are in the same form as matrices which put them in the same category.

Operator in mathematics is just a short hand for summarizing big data. It is easier to remember a straight line fit with two parameters than thousands of points in the input data. Operator like wave equation is a good short hand with a few symbols in the equation. Does this short hand describe all data, especially real data in the field? We have to face the reality of using both short hand and collecting massive real data. The residual is what we cannot explain with the operator. But the residual of real data is not necessarily random noise and the residual is quite interpretable.

The co-existence of human interpretation of hard equations and machine learning with soft equations might be the future of mathematics. “Abstract” topics like topology and semigroup should be encouraged as well as “practical” applications like neural network and topological data analysis. We have rushed into applications and have forgotten about significant old connections. Semigroup is a generalization of group theory which has many applications for a century. We did not venture further into semigroup theory because it is “messy” and does not have “clean” structural theorems as in group theory. However, semigroup or some other generalization is needed in dealing with real data which is predominantly “lossy” signal. Semigroup has had connection to finite-state automaton which is a mathematical abstraction of a physical computer.

Operator converts input data into output data. Operator could be pre-determined or post-determined. Pre-determined operator like differential equation is decided before seeing the data. Post-determined operator is decided after we see the data. It requires an extra step of data analysis before we decide on the operator in our toolbox.

We need both quantitative and qualitative methods to solve real world problems. Quantitative method involves "hard" equations like differential equation. Qualitative method involves interpretation and "soft" equations like topology, semigroup, AI, etc. It is an anxious time when there is no well-defined path. It is also an exciting time… perhaps just to have a glimpse of the mind of God.