Malcolm Adams

Luke Allen

Methods of Quantum Mechanical Machine Learning

Quantum Mechanics is the theory that describes matter at the atomic level. Atomic systems defy modeling by conventional methods due to the principals of superposition and uncertainty. As such, quantum mechanical systems require a different mathematical approach than the conventionally understood theory of Classical Mechanics. The traditional theory of mechanics sought to model systems by the Euler-LaGrange equations of motion. Differential equations derived from the Hamiltonian of the system in question. By knowing the degrees-of-freedom for a system, the equations of motion provide a framework for all possible states that a system may occupy. While the complete modeling of a system is only possible when there is a but a single degree-of-freedom, in the case of classical mechanics, the addition of multiple degrees of freedom does not overly complicate the mathematical description. It does not alter the Hamiltonian, nor does it "erase" any information previously known about the system. It simply increases the number of equations to be solved for. The calculations themselves are not made exponentially more difficult. Tragically, such is not the case for quantum mechanical systems. In quantum states, additional degrees of freedom mean that the system itself is physically more difficult to describe, regardless of whether or not the system has actually shifted. This is because describing a particle requires describing it as a superposition of all it's possible states. The states cannot be isolated, solved individually, and recombined to model the system.

This makes solving quantum systems for anything larger than a hydrogen atom, very computationally expensive. In the words of Paul Dirac himself: "It is desirable that approximate practical methods of quantum mechanics be developed, which can lead to an explanation of complex atomic systems without too much computation". [1] The prevailing methods used for quantum systems in computational physics are Hartree-Frock, and Density Functional Theory. Both of which seek to solve the Schrodinger Equation by "approximation" rather than explicit calculation. In the case of the Hartree-Frock theory this is done by attempting to solve the wave function for smaller degrees-of-freedom and using those solutions to approximate the wave function in the context of all degrees-of-freedom present. For Density Functional Theorem, the idea is to frame the wave function using a functional of electron density and energy functions. [2, 3]These methods have proven to be successful in recent decades, but as Dirac warned, they are still computationally time consuming. The use of these methods for even basic systems can still lead to calculations that can take on the order of days to finish. This is where Quantum Mechanical Machine Learning (abbr. QMML) comes into play, with the goal of not necessarily solving a perfect model, but to provide a close approximation.

One common problem for which QMML is quite helpful is in chemistry. The problem, related to 3-D modeling, is to find the enthalpy of atomization of a material. To calculate the atomization energy for some molecule, one must know the cumulative energy present in every bond of a molecule as well as how every single atom present is affected by every other. A daunting task. One that is solvable, but not in any kind of efficient manner. Problems like this are precisely what Dirac felt would be better served by "Quantum Approximation." Enter: Neural Nets

QMML is aimed at solving the energy eigenvalue equation by using a neural net. Neural Networks are fantastic function approximators. By using neural nets to find probable solutions for things

like enthalpy of atomization, computational costs can be minimized dramatically. What once took days or even weeks to calculate, could be reduced to as low as a few seconds. Machine learning presents an entirely different approach to scientific problem solving that could prove utterly vital in our understanding of physical systems.

For the question of atomization energy, the atomic construction and coulomb matrix are two primary ways of representing a molecule such that the data can be put into neural network. [4] The atomic construction for a compound of interest begins with representing the possible molecules present in a material as a vector:

$$\Omega = \{\Omega_1, \Omega_2, \Omega_3, \ldots, \Omega_m\} \tag{1}$$

Where each Ω_m gives a molecule that is present in the compound. Each Ω_m then is itself a combination of some number of atoms of a certain type. This leads to the definition of the atomic space:

$$A = \{A^1, A^2, A^3, \ldots, A^K\} \tag{2}$$

Where each A^k gives a particular atom known to be present in the compound. Using this, each molecule Ω_m, can be presented simply by its chemical formula. With the integer numbers of each atom A^k that compose it. That is,

$$\Omega_m = \alpha_1 A^1 \, \alpha_2 A^2 \, \alpha_3 A^3 \ldots \, \alpha_k A^k \tag{3}$$

Where α_k gives the number of the atom A^k present in the molecule. If the atomic space is known, that is, if the elements of A are known to be bounded above by K, then the specific atoms need only be indexed by integers [1, K]. And the atomic composition of a compound can then be given as an M x K matrix:

$$X = \begin{bmatrix} \Omega_1 \\ \Omega_2 \\ \Omega_3 \\ \vdots \\ \Omega_m \end{bmatrix} = \begin{bmatrix} \alpha_1^1 & \alpha_1^2 & \alpha_1^3 & \ldots & \alpha_1^k \\ \alpha_2^1 & \alpha_2^2 & \ldots & \ldots & \alpha_2^k \\ \vdots & & & & \\ \vdots & & & & \\ \alpha_m^1 & \ldots & \ldots & \ldots & \alpha_m^k \end{bmatrix} \tag{4}$$

An example Atomic composition, using the space A = { H, O, C }, (hydrogen, Oxygen, Carbon) is given as follows:

$$X = \begin{bmatrix} H_2 \\ O_2 \\ H_2O \\ CH_4 \\ CO_2 \\ \vdots \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 2 & 1 & 0 \\ 4 & 0 & 1 \\ 0 & 2 & 1 \\ \vdots & \vdots & \vdots \end{bmatrix} \tag{5}$$

The big problem however, with the atomic composition matrix is that simply knowing the amount of each type of element in a molecule is not enough to describe it uniquely. There are numerous materials, specifically isomers, that may share the same chemical formula between two different compounds. Having the same atomic makeup does not guarantee that the molecular structure will be identical. [4]

In addition to the Atomic composition matrix, there is the Coulomb matrix. Which attempts to encode the potential energy relationships between each pair of charges in a given molecule. In doing so, the Coulomb matrix effectively replaces the potential energy function in the Schrodinger equation. The elements of the coulomb matrix itself consist of the pairwise potential energies for each pair of electrons in the molecule. The elements are given by coulombs law as follows:

$$C_{ij} = \begin{cases} \frac{Z_i Z_j}{r_{ij}} & i \neq j \\ \frac{1}{2} Z_i^{2.4} & i = j \end{cases} \tag{6}$$

With Z_i being the charge of atom i, and r being the distance in between the atoms i, and j. [1, 4] The coulomb matrix is invariant under rotation and translation, which is convenient, though it is not without problems. Specifically, there is no way to objectively order the matrix in terms of the atoms present in a molecule. The same molecule can be described by M! different matrices. [1] Another issue in the context of Neural Networks, is that depending on the material in question, the coulomb matrix may be of a varying size. This does not work for machine learning and requires that matrices either be padded with 0 elements on the outside in order to fit an input for the network, or that a different neural network be used all together for different molecules.

In the context of quantum mechanics, a very significant aspect of the coulomb matrix is that is allows for the re-forming of the Hamiltonian operator H in the Schrödinger equation:

$$H\Psi = E\Psi$$

As follows, [1]

$$(H_{electron} + H_{nucleus} + \frac{1}{2}\sum_{ij} C_{ij})\Psi = E\Psi \tag{7}$$

This has separated the Hamiltonian operator from the traditional kinetic and potential energies into the energies of the electrons and the nuclei of the molecule. This formulation eases the computational load of Hartree-Frock and Density functional theorem mentioned above. And allows a deep learning algorithm to apply these theorems in a manner that is orders of magnitude faster computationally than was previously achievable.

Circling back, we now have what is needed to find the atomization energy of any compound by using QMML. Both the atomic composition and coulomb matrices have been used to great degree of success in modeling the electronic properties of molecules. The input for the network is to be of the form of one of the matrices described above. This gives the network all the information that would be significant to determine the energy of the molecules present. The output of the network is a vector:

$$Y = [y_1, y_2, y_3, \ldots, y_m] \tag{8}$$

Where each component y_m is a real number that corresponds to the atomization energy of the m^{th} molecule present in the compound. [4] The deployment of neural networks in this manner allows for a remarkably efficient means of approximating the electronic density of molecules vastly more complex than that of a simple hydrogen atom. The output of the network is a solution to the electronic Schrödinger equation. What remains becomes a problem of optimization; How to refine the input and train the network to produce the most accurate approximation possible.

For optimizing QMML in regard to the Enthalpy of Atomization problem, the first roadblock comes in representing a molecule in a way that is entirely unique. The coulomb matrix in particular, is subject to many different possible permutations for the same molecule. The way that this can be remedied is by examining the L2-norms of each permutation and ordering them in descending order. The matrices are symmetric, and so only their lower triangular part need be considered. Using this, each of the m-rows would correspond to the coulomb matrix values of a particular molecule, and since the matrices were put in descending order, the m-molecules would also be in descending order by number of atoms. [4] This would give a new input for the neural network in the form of an M x L matrix:

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1L} \\ x_{21} & x_{22} & \dots & x_{2L} \\ \vdots & \vdots & \vdots & \vdots \\ x_{m1} & \dots & \dots & x_{ML} \end{bmatrix} \tag{9}$$

Where the M-th row, represents the Coulomb Matrix values for the atoms in the M-th molecule.

Looking now at all the possible inputs, it is found that applying a combination of both the atomic composition matrix and the coulomb matrix re-formulated as above, yields an input that minimizes the possibility of a molecule being mis-identified by the network.

When seeking to optimize a neural network it is also of interest to determine the best way to train the network. As used by Tchagang and Valdés, a Bayesian method allows for a considerably more optimal treatment of data in QMML. A Bayesian problem solving method is used specifically because it seeks to find solutions that don't overfit the data, potentially finding problems that were never there to begin with. [4] For QMML, where the entire goal is to make computationally intensive problems less intensive, adopting a problem-solving strategy that is based in reducing the multiplicity of solutions seems quite imperative.

The use of neural networks to solve emerging problems has developed quite a bit in previous decades. It doesn't seem as though the theory of quantum mechanics itself will be made easier to apply anytime soon. Indeed, it certainly seems as if the future of science depends quite heavily on the further development and optimization of machine learning. It is unfortunate however, that the best we can do is simply attempt to find a good approximation. But for those who study quantum mechanical systems, the idea of being tasked with unconventional problems is nothing new. And in modern decades we have begun to develop an equally unconventional method for finding solutions to these distastefully complex problems. QMML is an incredibly exciting field when one considers the volume of systems thought by many brilliant minds to be unsolvable, where answers would only be found in the form of loose approximations at best.

Works Referenced:

[1] Laksh. “Teaching AI How to Do Quantum Mechanics.” *Medium*, Towards Data Science, 6 Aug. 2019, towardsdatascience.com/teaching-ai-how-to-do-quantum-mechanics-8ea693b2fa8a.

[2] Sherrill, David. *An Introduction to Hartree-Fock Molecular Orbital Theory*. June 2000, vergil.chemistry.gatech.edu/notes/hf-intro/hf-intro.pdf.

[3] *Density Functional Theory for Beginners*, newton.ex.ac.uk/research/qsystems/people/coomer/dft_intro.html.

[4] Tchagang , Alain B, and Julio J Valdés. *Prediction of the Atomization Energy of Molecules Using Coulomb Matrix and Atomic Composition in a Bayesian Regularized Neural Networks*. arxiv.org/ftp/arxiv/papers/1904/1904.10321.pdf.