Variational quantum eigensolvers

A hybrid classical-quantum optimization algorithm

Ron K. Cytron and students (to be listed)

Washington University
Saint Louis, Missouri

March 27, 2024

Overview

A brief introduction

1. Given a matrix (H), we wish to find the set of eigenvalues associated with H. VQE's goal is finding the smallest eigenvalue of H, but it can be modified to also find higher-level eigenvalues
2. We can analytically find the eigenvalues of H through matrix decomposition methods. The issue arises when H becomes large enough that it can no longer be analytically solved.
3. As of 20 years ago, the largest classically solved matrix was through Google's PageRank algorithm. The generated eigenvector was 2.7 billion web pages long.

Brief Mathematical Recap

Quantum Kickstart

Eigenvalue Equation:

$$H|\Psi\rangle = E_n|\Psi\rangle$$

Definitions

1. H - the Hamiltonian Matrix. In physics, this represents the total energy, expressed as $H = KE + PE$, where KE is the kinetic energy and PE is the potential energy. It can also represent optimization problems and other complex systems. Generally, it is the large matrix we are attempting to solve.
2. $|\psi\rangle$ - a State Vector in quantum mechanics, commonly referred to as a "ket". It is a general eigenvector of the Hamiltonian Matrix.
3. E_n - A discrete energy level assuming H is the total energy. In mathematical terms, this is the eigenvalues of H. In VQE, we are searching for E_0.
4. $|x\rangle$ - a state vector of the computational basis. For 1 qubit, $|0\rangle$ and $|1\rangle$.

VQE Pipeline

How the algorithm actually works

The Steps of VQE given step by step

1. Hamiltonian Construction and Representation
2. Encoding of the Operators
3. Ansatz and State Preparation
4. Parameter Optimization

VQE Pipeline

An Overview

1. Hamiltonian Construction and Representation
2. Encoding of the Operators
3. Ansatz and State Preparation
4. Parameter Optimization

The blue steps are done using classic computational methods either using a classical computer or analytical methods

The green step is done using a quantum computer

Hamiltonian Assumptions

Why is this faster anyway?

1. In the long term, Quantum Computers will offer speedup for full analytical solutions of Hamiltonian matrices (FQE), but this will require millions of physical qubits for actualization
2. For VQE, we will assume the complex Hamiltonian is actualized and defined by the system (Step 1)
3. For any complex Hamiltonian, we also assume it can be decomposed into the addition of Pauli Strings $(P_i \otimes P_i \otimes ... \otimes P_i) \quad P_i \in \{X, Y, Z, I\}$ (Step 2)
4. For a NxN matrix, we will have at most O(Poly(N)) Pauli Strings for an NxN Hamiltonian. The key is utilizing a decomposition method which limits the size of our Pauli Strings.
5. We will not go into these decomposition methods analytically. Instead, we will assume that any unitary H can be decomposed into these strings. For more information, an example is the Jordan-Wigner Decomposition

Measuring with the Hamiltonian- Pauli Gate Tricks

On a Quantum Computer

We wish to measure in the Z basis to receive results akin to $|x\rangle$

Measurement can only be done in the Z basis, but by rotating the circuit you can emulate measurement in that basis.

$\mathbf{X}$: $\mathbf{H}$ Hadamard

$\mathbf{Y}$: $R_x(\frac{\pi}{2})$

$\mathbf{Z}$: no rotation

$\mathbf{I}$: no measurement performed

*note expectation values of the Pauli operators map a measurement of $|0\rangle$ to 1 and $|1\rangle$ to -1 due to the eigenvalues of the Pauli Z gate.

Pauli Strings

Representation and Measurement

For instances in which the Hamiltonian consists of terms with more than one Pauli operators multiplied together, the terms can be simplified to a single Pauli operator. This is elements of the change of basis operators simplified. The following table shows the relationships:

multiply	$\mathbf{X}$	$\mathbf{Y}$	$\mathbf{Z}$
$\mathbf{X}$	$\mathbf{I}$	i$\mathbf{Z}$	-i$\mathbf{Y}$
$\mathbf{Y}$	-i$\mathbf{Z}$	$\mathbf{I}$	i$\mathbf{X}$
$\mathbf{Z}$	i$\mathbf{Y}$	-i$\mathbf{X}$	$\mathbf{I}$

For example take the Hamiltonian for a single qubit:

$$H = \alpha_1 \mathbf{XX} + \alpha_2 \mathbf{YZ} + \alpha_3 \mathbf{IZ}$$

This can be simplified to the following where i is just a phase factor and can be ignored: $H = \alpha_1 \mathbf{I} - \alpha_2 i\mathbf{X} + \alpha_3 \mathbf{Z}$

Tricks we'll use!

- Generalized State Vector: $|\psi\rangle = \sum_x \alpha_x |x\rangle$
- Selection Vector: $\langle x|\psi\rangle = \langle x| \sum_y \alpha_y |y\rangle = \sum_y \langle x|y\rangle = \alpha_x \langle x|x\rangle = \alpha_x$
- Probabilistic Interpretation: $P(|x\rangle) = |\alpha_x|^2 = |\langle x|\psi\rangle|^2$
- Classical Expectation Value: $E(x) = \sum_x xP(x)$
- Quantum Expectation Value: $\langle H\rangle_\psi = \langle\psi|H|\psi\rangle$

Expectation Values

What did you expect?

- Proving that our quantum expectation value formulation is equal to the classical definition

$$
\begin{aligned}
&\langle\psi|H|\psi\rangle \\
&\qquad = \sum_{x,y} \alpha_y^* \alpha_x \langle y|H|x\rangle \\
&\qquad = \sum_{x,y} \alpha_y^* \alpha_x E_n \langle y|x\rangle \\
&\qquad = \sum_{x} \alpha_x^* \alpha_x E_n \\
&\qquad = \sum |\alpha_x|^2 E_n \\
&\qquad = E(E_n)
\end{aligned}
$$

Variational Principle

A Simple Bound

- Assume that $|\psi\rangle$ is no longer a general statevector, and instead a specific vector which we sample, $|\psi_n\rangle$
- All returned eigenvalues are guaranteed to be above the lowest eigenvalue in the spectrum, E_0, intuitively.
- Mathematically, we declare the effects of this as the variational principle:

$$\langle H \rangle_{\Psi_n} = \langle \Psi_n | H | \Psi_n \rangle = E_n \langle \Psi_n | \Psi_n \rangle \geq E_0 \langle \Psi_n | \Psi_n \rangle$$

- This formulation allows us to morph the eigenvalue problem into a minimization problem. We are seeking the minimum value of E_0 by calculating expectation values. But how do we vary our expectation value?

A Variable $|\psi\rangle$

The Ansatz

- Visualize our vector $|\psi\rangle$ as a generic point on the Bloch Sphere
- There is a point, or specific vector $|\psi_n\rangle$, which achieves our minimum eigenvalue E_0
- Our goal is parameterizing a given $|\psi_n\rangle$ to achieve that minimum point
- This variable statevector is called **The Ansatz**
- An ansatz which spans the whole Bloch Sphere is preferred, but can lead to problem slowdown. Ansatzs which only cover specific regions can also be used if the type of solution is known.

Steps for a Quantum Computer Scientist

- Prepare a Hamiltonian which represents the system and decompose it into Pauli Strings
- Prepare a parameterized wavefunction with a set of parameters **x**
- Apply $|\psi_n\rangle$ and H to formulate an Expectation Value on a quantum computer. (See Qiskit Estimators). Rotation may be neccessary to get back into the Z basis.
- Store the resulting eigenvalue: E_n and parameters: **x** in a classical dictionary
- The classical optimizer will use **x** and E_n to minimize the results of the expectation value.

A Simple Example

Step 1 and 2 of the pipeline

Let's say that hypothetically, we have the simplest system imaginable. Even simpler than the hydrogen atom as even the Hamiltonian for just a proton and an electron is egregious.

Step 1: Scientists figure out the Hamiltonian for this system looks like the following written in the Hamiltonian basis:

$$H = 3\,|0\rangle\langle 0| + 1\,|0\rangle\langle 1|$$

Step 2:When written this way the Hamiltonian encodes easily to the Pauli basis:

$|0\rangle\langle 0| = 1/2(\mathbf{I} + \mathbf{Z})$
$|0\rangle\langle 1| = 1/2(\mathbf{X} + i\mathbf{Y})$
$|1\rangle\langle 0| = 1/2(\mathbf{I} - \mathbf{Z})$
$|1\rangle\langle 1| = 1/2(\mathbf{X} - i\mathbf{Y})$

$$H = 1.5\mathbf{I} + 1.5\mathbf{Z} + 0.5\mathbf{X} + 0.5i\mathbf{Y}$$

Continuing the Simple Example

Preparing the Ansatz

- Our Ansatz: $|\psi_n\rangle = R_y(\theta)|0\rangle$ (simple but could not span all possible solutions)
- Our parameter to be optimized ($\mathbf{x}$): θ
- Hamiltonian: $H = 1.5\mathbf{I} + 1.5\mathbf{Z} + 0.5\mathbf{X} + 0.5i\mathbf{Y}$
- Since a measurement is only an approximation of the expectation value, let's say that we want to do 100 shots of each run.
- We prepare 300 circuits. 100 with only our Ansatz to represent $\mathbf{Z}$ in our Hamiltonian. 100 with our Ansatz and a Hadamard to represent $\mathbf{X}$ in our Hamiltonian. 100 with our Ansatz and $R_x(\frac{\pi}{2})$ ($\mathbf{Y}$). $\mathbf{I}$ can be ignored and can be added back later

The Circuit for the Example

visualizing the implementation

Let's being with an initial guess of $\theta = \pi/2$.
An example of what one of the prepared circuit may look as follows. Since our Hamiltonian only involves one qubit we can prepare the different measurements together as they don't interact with one another:

q[0] is our measurement in the $\mathbf{Z}$ basis. q[1] is our measurement in the $\mathbf{X}$ basis, and q[2] is our measurement is the $\mathbf{Y}$ basis. The R_y gate is our ansatz and θ is set to $\frac{\pi}{2}$. We prepare 100 of the above circuits. Whenever we measure q[0] as $|0\rangle$ we add -1 to our calculation for the Hamiltonian expectation value in the $\mathbf{Z}$ basis. If we measure $|1\rangle$ we add 1. We repeat for each quibit.

Finding the Expectation Value

for the simple example

We are solving

$\langle H \rangle_{\Psi(\frac{\pi}{2})} = 1.5\langle\Psi(\frac{\pi}{2})|\mathbf{I}|\Psi(\frac{\pi}{2})\rangle + 1.5\langle\Psi(\frac{\pi}{2})|\mathbf{Z}|\Psi(\frac{\pi}{2})\rangle + 0.5\langle\Psi(\frac{\pi}{2})|\mathbf{X}|\Psi(\frac{\pi}{2})\rangle$

$+0.5\langle\Psi(\frac{\pi}{2})|i\mathbf{Y}|\Psi(\frac{\pi}{2})\rangle$

We do not need to perform any measurement for *Identity*, so the expectation value is just one. $\langle\Psi(\frac{\pi}{2})|\mathbf{I}|\Psi(\frac{\pi}{2})\rangle = 1$

For the $\mathbf{Z}$ measurement

$\langle\Psi(\frac{\pi}{2})|\mathbf{Z}|\Psi(\frac{\pi}{2})\rangle = (1+1-1+1...+1-1)/100 = \frac{-40+60}{100} = 0.2$

For the $\mathbf{X}$ measurement

$\langle\Psi(\frac{\pi}{2})|\mathbf{X}|\Psi(\frac{\pi}{2})\rangle = (-1-1-1-1...-1-1)/100 = \frac{-100}{100} = -1$

For the $\mathbf{Y}$ measurement

$\langle\Psi(\frac{\pi}{2})|\mathbf{Y}|\Psi(\frac{\pi}{2})\rangle = (1+1-1+1...+1-1)/100 = \frac{-43+57}{100} = 0.14$

Finding the Expectation Value Cont.

Now we just plug the measured values into the expectation value equation.

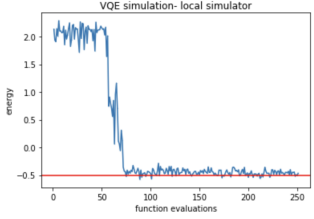
$$\langle H \rangle_{\Psi(\frac{\pi}{2})} = 1.5 * 1 + 1.5 * 0.2 + 0.5 * (-1) + 0.5 * 0.14$$

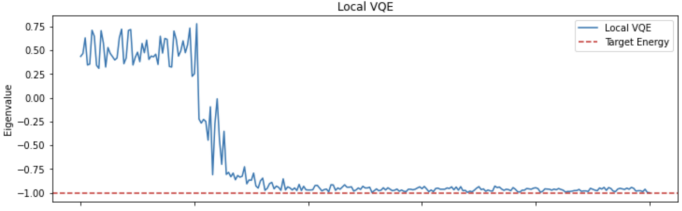
$$\langle H \rangle_{\Psi(\frac{\pi}{2})} = 1.5 + 0.3 - 0.5 + 0.07 = 1.37$$

We then store the expectation value of 1.37 with a θ value of $\pi/2$. This process will be repeated given a new guessed θ by the classical optimizer

Example VQE Runs

Notice that the energy is dipping below the minimum value. How is this possible?





Optimization Issues

The optimizers used for VQE are classical processes. Trials runs observed clear values below the ground state energy, which should not be possible due to the variational principle. Previously, we assumed the following definition of the Hamiltonian:

$$H = \sum_{i}^{n} \alpha_i P_i$$

.

The typical variance relationship in quantum mechanics is:

$$< (\Delta H)^2 > = < H^2 > - < H >^2$$

Although the Hamiltonian can be decomposed into Pauli operators, its variance does not decompose as smoothly, causing an error factor to become present

Hamiltonian Variance

$$\langle H \rangle_\Psi = \alpha_1 \langle \Psi_n | P_1 | \Psi_n \rangle + \alpha_2 \langle \Psi_n | P_2 | \Psi_n \rangle + ... + \alpha_n \langle \Psi_n | P_n | \Psi_n \rangle$$

$$\langle H^2 \rangle_\Psi = \alpha_1^2 \langle \Psi_n | P_1^2 | \Psi_n \rangle + \alpha_2^2 \langle \Psi_n | P_2^2 | \Psi_n \rangle + ... + \alpha_n^2 \langle \Psi_n | P_n^2 | \Psi_n \rangle$$

The Hamiltonian, as an operator, has definite states. This means typically,

$$< (\Delta H)^2 >= 0$$

It is clear that is not the case here, given our decomposition of $< H >$ and $< H^2 >$. The terms will not cancel out, forming a nonzero variacnce relation. Since the Hamiltonian is not a determinate operator, similar to how we consider it in quantum mechanics, we cannot consider every measurement of energy to be purely definite. Some variability, or error, will exist.

Proposed Error

This is taken from a paper by Kandala, Mezzacapo, et al written in 2017 about using VQE for the modeling of small molecules and quantum magnets. Let S be the number of samples, T be the length of the Pauli Strings, and h_{max} be the max weight. The variance can be defined:

$$Var[H] = \sum_{\alpha}^{T} h_{\alpha}^2 < \Delta P_{\alpha}^2 >$$

$$\epsilon = \sqrt{\frac{Var[H]}{S}} \leq \sqrt{\frac{T|h_{max}|^2}{S}}$$

With this error, although a value below the ground usually cannot be produced by the Hamiltonian, it possible to just break over the cusp. Different methods for Pauli decomposition and different optimizers can have different effects on this issue.