STOCHASTIC GRADIENT DESCENT FOR HYBRID QUANTUM-CLASSICAL OPTIMIZATION

Rajesh Sathya Kumar

***Abstract*—Stochastic Gradient Descent is a common algorithm for optimization in classical computing. In this paper, we discuss about a similar algorithm which uses the expectation values of finite number of measurement samples to form the Quantum version of Stochastic Gradient Descent (QSGD) [1]. We will discuss and implement Single Shot Gradient Descent, Doubly Stochastic Gradient Descent and Adaptive Stochastic Gradient Descent algorithms. The convergence properties of these algorithms will be discussed and compared using the experimental results. It is shown that Quantum Stochastic Gradient Descent algorithms have low convergence rate than other Quantum Optimizers, but faster iterations.**

Keywords—Quantum Computing, Stochastic Gradient Descent, Convergence Properties

I. Project Description

Quantum Computing has shown a lot of promise is having a computational advantage over the classical counterparts. One of the most obvious applications is Optimization and Stochastic Gradient Descent is one of the most common optimization algorithms in classical computing. The algorithm has many applications in Machine Learning, Signal Analysis, Compression, Audio etc. In this paper, we will study about the quantum implementations of stochastic implementations and their convergence properties. Other variations of GD are Quantum Natural Gradient (QNG), which is a direct application of Variational Quantum Algorithm (VQA) and individual Coupled Adaptive Number of Shots (iCANS) optimizer which can adapt a shot number of measurements [6] [4] [3] [2]. In general, Stochastic Gradient Descent optimizers have the following properties. They typically can estimate the gradient faster with lesser number of random samples, the stochasticity helps to avoid local minima and saddle points and better convergence properties than regular gradient descent [5]. In SGD, the parameter update rule is given by,

$$\theta(t)= \theta(t) - \eta\, \nabla L\big(\theta(t)\big)$$

To generalize, it can be modified as below,

$$\theta(t)= \theta(t)- \eta\, g^{(t)}\big(\theta(t)\big)$$

where η is the step size and $g^{(t)}(\theta)$ is a random variable such that,

$$E\big[g^{(t)}\big(\theta(t)\big)\big]= \nabla L(\theta)$$

In Variational Quantum Algorithms (VQA), we optimize the quantum circuit $U(\theta)$ to minimize the Hamiltonian or the expectation value of the circuit [3]. The expectation of the circuit is given by,

$$\langle A_i \rangle= \langle 0|U(\theta)^* A_i U(\theta)|0\rangle$$

where $\{A_i\}$ are the list of observables and $L(\theta, \langle A_i\rangle, \dots, \langle A_M\rangle)$ are the loss functions.

The aim of this paper is to implement the stochastic gradient decent using a Quantum python SDK, Pennylane. It is a cross-platform python library for differentiable programming with Quantum Computing. We will use matplotlib to draw the convergence curves and execution times of each of those algorithms and discuss the different convergence properties. The project implementation contains three parts. First, we discuss and implement the Single-Shot Stochastic Gradient Descent for Variational Quantum Eigensolver (VQE) [1]. The convergence curve of vanilla gradient descent and doubly Stochastic Gradient Descent is given below.

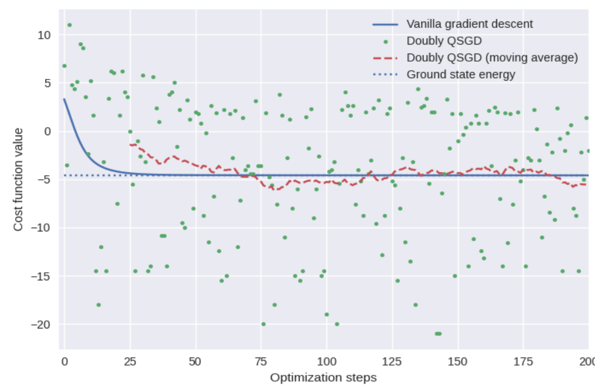

Figure 1: Convergence Curve of QSGD and Vanilla GD

II. Project Plan

The tentative project plan for the implementation of the algorithm is given below.

Algorithm	Tentative Completion Date
Single Shot SGD	Oct 11, 2022
Doubly SGD	Oct 25, 2022
Adaptive Stochasticity	Nov 8, 2022

References

[1] Ryan Sweke, Frederik Wilde, Johannes Jakob Meyer, Maria Schuld, Paul K. Fährmann, Barthélémy Meynard-Piganeau, Jens Eisert. "Stochastic gradient descent for hybrid quantum-classical optimization." arXiv:1910.01155, 2019.

[2] Kübler, J. M., Arrasmith, A., Cincio, L., & Coles, P. J. (2019). An Adaptive Optimizer for Measurement-Frugal Variational Algorithms. arXiv. https://doi.org/10.22331/q-2020-05-11-263

[3] Schuld, M., Bergholm, V., Gogolin, C., Izaac, J., & Killoran, N. (2018). Evaluating analytic gradients on quantum hardware. arXiv. https://doi.org/10.1103/PhysRevA.99.032331

[4] Shun-Ichi Amari. “Natural gradient works efficiently in learning.” Neural computation 10.2, 251-276, 1998. James Stokes, Josh Izaac, Nathan Killoran, Giuseppe Carleo. “Quantum Natural Gradient.” arXiv:1909.02108, 2019.

[5] Aram Harrow and John Napp. “Low-depth gradient measurements can improve convergence in variational hybrid quantum-classical algorithms.” arXiv:1901.05374, 2019.

[6] Naoki Yamamoto. “On the natural gradient for variational quantum eigensolver.” arXiv:1909.05074, 2019.