Exploring Quantum Linear Regression

Victor Shyaka [1]

[1] Computer Science, Stanford

Introduction

Quantum computing, particularly in machine learning, offers a substantial speedup over classical methods. Quantum Machine Learning (QML) integrates classical and quantum techniques, with Quantum Linear Regression (QLR) as a notable example. In QLR, classical data is encoded into qubits for processing by quantum algorithms. While classical Linear Regression using the Normal Equation has a computational complexity of $O(nm^2 + m^3)$ for n samples and m features, QLR operates at a complexity of $O(\log(nm)s^3\kappa^6\epsilon^{-1})$, where s is the matrix sparsity, κ the condition number, and ϵ the accuracy level. Significant research has focused on minimizing runtime dependencies in Quantum Linear Regression (QLR), yet this study endeavors to demonstrate optimization potential without modifying the QLR algorithm's intrinsic structure. This is achieved by linearly modifying a dataset and regularizing through both classical and quantum linear regression models.

$$s \propto \frac{\#\text{ Zero elements}}{\#\text{ Non-zero elements}}\ ;\ \kappa \propto \frac{\text{Largest element}}{\text{Smallest element}}$$

Data, Features

Our training data was randomly generated based on a housing price problem and was set to 16 samples with 4 features, namely: square footage, number of bedrooms, number of bathrooms and a bias feature. Here are 3 samples from our dataset:

Sq.Ft	Bed	Bath	Price
4234.5	3	3	6746.1
3280.1	4	3	6182.0
3685.1	1	1	4850.3

Table 1. Data without modification.

We use 16 data samples and 4 features because qubit systems use binary arithmetic and therefore require a space that's a power of 2 for encoding. A smaller dataset also involved shorter runtimes and would not hinder a comparative analysis. Our linear modification to the dataset transforms and features with a factor of 1000.

Sq.Ft	Bed	Bath	Price
4.2345	3	3	6.7461
3.2801	4	3	6.1820
3.6851	1	1	4.8503

Table 2. Data with modification.

Our testing data was of sample size 4, and was modified like our training data.

Model

- **Classical LR model**: Our linear model solved the following linear equation: $X^TX\theta = X^Ty$, where X is the training data matrix, y is the label vector and θ are the parameters we are trying to find. We then generate predictions by $\tilde{X}\theta$, where $\tilde{X}$ is the testing data. This is the regular normal equation model for linear regression, no modifications were performed to it.
- **Quantum LR model**: The linear model can be split into **amplitude encoding** and solving the linear equation through the **Harrow-Hassidim-Lloyd algorithm (HHL)**. First the training data, X, is encoded as follows: $X = \frac{1}{\sqrt{\ell^2\text{-norm}}}\sum_{j=0}^{n-1}|j\rangle$, where $|j\rangle$ is our qubit state and m is the number of samples. The HHL algorithm then proceeds to compute the θ parameter by computing relevant eigenvalues and inverses of our aforementioned linear equation. Our modification involves setting up **Tikhonov Regularization** of the form $X + I\lambda$, for various small positive λ values before the HHL algorithm.

Results

On un-modified data, our CLR model gave a Mean-Squared Error of **0.007** on predictions. With the same data our QLR's κ factor was of order 10^5. When modified, the QLR model gave a κ factor of order 10^2 and a Mean-Squared Error of **0.026** on predictions. The regularization reduces the κ factor even further, allows for faster runtimes and improves on the MSE.

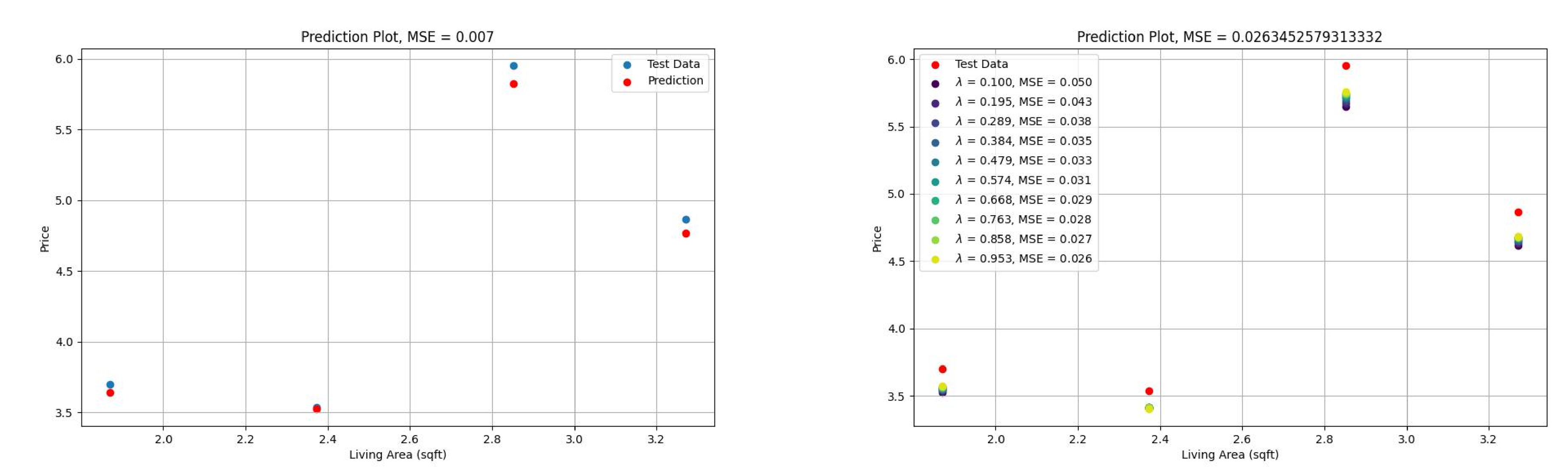

Figure 1. Left: CLR model predictions. Right: QLR model predictions.

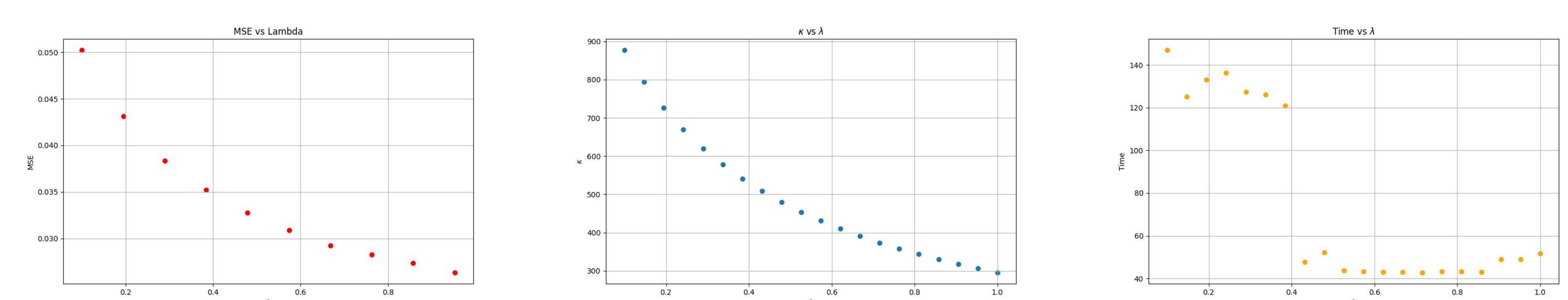

Figure 2. Left: Change in MSE based on λ. Middle: Change in the Condition Variable based on λ. Right: Change in Time based on λ.

Discussion

We observe an improvement in the MSE due to the Ridge Regularization or Tikhonov Regularization. This form of regularization has been shown to redefine ill-posed linear algebra problems. What was intriguing to observe are improvements in the κ factor due to linear data modification. These modifications had the desired effect because the breadth of the values was cut by our reduction factor. From how κ is defined, the largest and smallest element ratios were subsequently shrunk by the same amount. One aspect that surprised us was the time distribution, which suggested that there were other factors that we did not account for in our analysis, they might have been other parameters such as sparsity, or the over fitting and under fitting problem. Another indication of this is the MSE comparison between our two models. We expected the results to be closer in MSE, because we assumed the QLR would perform similarly to the CLR, however the differences in their model construction could also be playing a major role in output.

Future Work

The next steps for this research would be first to vary sample sizes. We would expect the CLR to give better predictions and it would be interesting to monitor the QLR's behavior. It would also be interesting to explore the role over-fitting plays in QLR. Given our sample size to feature ratio and how correlated our features were, both our models visibly over-fitted to noise in our data. A keen analysis of over-fitting over other data sizes would be interesting to pursue.

References

[1] M. Schuld , F. Petruccione "Supervised Learning with Quantum Computers" 2018.

[2] A. Zeguendry, Z. Jarir, M. Quafafou "Quantum Machine Learning: A Review and Case Studies" 2022.

[3] A. W. Harrow, A. Hassidim, and S. Lloys "Quantum Algorithm for Linear Systems of Equations" 2009.

[4] M. Schuld , F. Petruccione "Prediction by linear regression on a quantum computer" University of KwaZulu-Natal, 2016.

[5] C. Shao, H. Xiang "Quantum Regularized Least Squares Solver with Parameter" 2018.

[6] R. Schaeffer LaTeX Poster Template.

victorks@stanford.edu