Quantum-enhanced Convex Regression

Juan Carlos Gonzalez-Espitia[1,2] and Jannis Born[1]

[1]IBM Research Europe - Zurich, Switzerland
[2]Politecnico di Milano, Dipartimento di Matematica - Milano, Italy

May 23, 2024

Extended abstract

Convex Regression (CR) is a type of structured regression that strives to fit the best curve $f : \mathbb{R}^n \to \mathbb{R}$ to a data set $\mathcal{D}$ under the constraint of convexity, that is, $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)\ \forall x, y \in \mathbb{R}^n$, and $\lambda \in [0, 1]$ [5]. CR is of great interest in finance and economics where convexity constraints appear naturally for value functions. Some remarkable examples include the Kantorovich problem with quadratic costs [15], the construction of utility functions [2], and risk analysis [24], among others. Nonetheless, CR has been adopted in more general contexts due to its relationship with dynamic programming [4], convex optimization [25], and reinforcement learning [12, 30, 11]. Despite its widespread utility, CR has high computational costs [27] and often faces stability issues – especially for frameworks relying on Neural Networks (NN) [3, 23]. In modern approaches the complexities for unconstrained CR is $O(n^5d^2/\varepsilon)$ while for convex Lipschitz regression it is $O(n^3d^{1.5}/\varepsilon + n^2d^{2.5}/\varepsilon + nd^3/\varepsilon)$[28, 27] which quickly becomes prohibitive for high dimensional problems or big data.

Quantum computing offers a new paradigm which rapidly advances in hardware and error correction [7]. Hopes are high for practical utility in machine learning (ML) [1, 13] and beyond [16]. While unconstrained linear regression has been studied in a multitude of ways relying on quantum [29, 26], this work is the first to explore the problem of quantum **convex** regression. The main challenge for proposing Quantum CR (QCR) algorithms is the difficulty of encoding geometrical constraints, like convexity, directly into the components of the quantum circuits (QC) because quantum gates are represented as unitary operators on a Hilbert space. Classical NNs may guarantee positivity of the network parameters and convexity of the activation functions [3] but there is no way to impose similar conditions in QCs emulating NNs. Recently, splines were used to model arbitrary non-linear activation functions as the solutions to linear systems [17, 18, 14]. This enables to formulate the regression problem as linear systems that can be solved efficiently via Quantum Linear Solvers thus connecting linear regression with Quantum Variational Circuits. In the long term, QCR might compete with classical CR and thus find application in robotics, reinforcement learning [9] or as an alternative to input-convex NNs [3] which are commonly utilized in neural optimal transport to exploit Brenier's theorem [8, 19, 20].

Here, we study convex interpolation (as a particular case of convex regression) and apply the parameterization by [21] to formulate a linear system for convex splines. We devise SWAM, a novel algorithm to set the boundary conditions of the splines curves, and solve the sparse linear system using the Variational Quantum Linear Solver (VQLS) [6] whose heuristic time complexity scales efficiently with ε, n, d, and the matrix condition number κ. Our method theoretically outperforms classical schemes as long as the underlying quantum states can be prepared efficiently. We expect that upcoming advances in efficient quantum state preparation will enable general utility of the proposed technique not only for convex interpolation, but also for convex regression.

Formulation

We define the single-segment spline problem in $\mathbb{R}$ as follows. Given a couple of points, $(x_0, s_0), (x_1, s_1) \in \mathbb{R}^2$, and $s'_0, s'_1 \in \mathbb{R}$, we want to find a curve $s : [x_0, x_1] \to \mathbb{R}$, $s \in C^2([x_0, x_1])$, such that $s(x_0) = x_0$, $s(x_1) = x_1$, $s'(x_0) = s'_0$, and $s'(x_1) = s'_1$. Defining $\xi(x) = \frac{x-x_0}{x_1-x_0}\ \forall\ x \in [x_0, x_1]$, we parameterize s as follows:

$$s(\xi) = \lambda^n \left(\frac{a}{n(n-1)} \left(\xi - \frac{1-\gamma}{2} \right)^n \qquad + \frac{b}{n(n-1)} \left(\xi - \frac{1+\gamma}{2} \right)^n \right) + c\xi + d,$$

where $a, b, c, d \in \mathbb{R}, 0 \leq \gamma \leq 1$, $n \geq 4$ are parameters that control the shape of the curve. After taking the first derivative of s, imposing the boundary conditions and performing some algebraic manipulations, we can write the problem as $A\zeta = \eta$, where $\zeta = [a, b, c, d]$ is the vector of unknowns, $\eta = [s_0, s_1, s'_0, s'_1]$, is the vector of parameters of the problem, and A is the following matrix:

$$A = \begin{bmatrix} \alpha_0 & \beta_0 & 0 & 1 \\ \beta_0 & \alpha_0 & 1 & 1 \\ \alpha_1 & -\beta_1 & 1 & 0 \\ \beta_1 & -\alpha_1 & 1 & 0 \end{bmatrix},$$

with $\alpha_0, \alpha_0, \beta_0, \beta_1$ constants defined in terms of the problem's parameters. For a multi-segment problem, instead of finding a single spline interpolating two points, we start with a set of "nodes" $X = \{x_0, x_1, x_2, \ldots, x_N\}$ and a set of "images" $S = \{s_0, s_1, s_2, \ldots, s_N\}$ that define N different single-segment spline problems with sub-domains in $[x_k, x_{k+1}]$ for $k = 0, 1, \ldots, N-1$. We want to find a function $f : [x_0, x_N] \to \mathbb{R}$ defined as the union of the N solutions of the single-spline interpolation problems. Additionally, we impose that the images and slopes of adjacent spline functions coincide when evaluated in the nodes. Because in the context of regression and interpolation nodes and their images are given, the problem is then reduced to find a valid set of slopes $\{s'_0, s'_1, s'_2, \ldots, s'_N\}$ such that they allow for the interpolation of the full set of points while preserving the global property of convexity.

The SWAM method and application of the VQLS

In the context of convex interpolation the set of slopes cannot be arbitrary because it determines the curvature of f and consequently has a direct impact the function's convexity. To set those values properly, we propose a modification to the Slope Averaging Method (SAM) [21]. We choose the segment's slopes as a weighted average of adjacent empirical slopes weighted by their relative distance. We call this method the Slopes Weighted Averaging Method (SWAM). Concretely, let us define $\Delta^+(k) = x_{k+1} - x_k,\ \Delta^-(k) = x_k - x_{k-1}$, and

$$\hat{s}^+(k) = \begin{cases} \frac{s_{k+1}-s_k}{\Delta^+(k)} & \text{if } \Delta^+ > 0 \\ 0 & \text{if } \Delta^+ = 0 \end{cases} \qquad \hat{s}^-(k) = \begin{cases} \frac{s_k-s_{k-1}}{\Delta^-(k)} & \text{if } \Delta^- > 0 \\ 0 & \text{if } \Delta^- = 0 \end{cases},$$

then, we select the slopes as:

$$\begin{cases} s'_0 = (1 - \sigma\ sign(\hat{s}^+_0))\ \hat{s}^+_0 \\ s'_k = \frac{\Delta^+(k)\ \hat{s}^+(k) + \Delta^-(k)\ \hat{s}^-(k)}{\Delta^-(k)+\Delta^+(k)}, \quad 1 \leq k \leq N-1 \\ s'_N = (1 + \sigma\ sign(\hat{s}^-_N))\ \hat{s}^-_N \end{cases}$$

While the SWAM method guarantees convexity whenever possible, a set of points that does not admit a convex interpolation curve will be interpolated in a non-convex manner by selecting the empirical slope. Since the problems inside each sub-domain are independent after setting the corresponding boundary conditions, we define an enlarged linear system $\mathcal{A}x = b$ containing the information of all segments and boundary conditions. The matrix $\mathcal{A}$ is a block diagonal matrix where the k-th diagonal element is the A_k matrix for the k-th sub domain; the vector x is the concatenation of the the vectors of unknowns $[a_k, b_k, c_k, d_k]$ for each sub-domain $k = 0, 1, \dots, N$; and the vector b is the concatenation of the vectors of parameters $[s_{0k}, s_{1k}, s'_{0k}, s'_{1k}]$. The new enlarged system is sparse of dimension $4N$ and possesses significant computational challenges due to the high complexity of the current available methods $O(n^5 d^2/\varepsilon) \approx O(N^7)$. As an attempt to ease the computational costs by leveraging on Quantum Computing techniques we use the Variational Quantum Linear Solver (VQLS) [6] to solve the extended linear system.

The goal of the VQLS is solve the linear system $Ax = b$ relying on Variational Quantum Circuits (VQCs). To implement it we prepare efficiently the quantum state $|b\rangle$ and a circuit encoding the classical information of $\mathcal{A}$. Then, we find the optimal set of parameters θ such that an ansatz $V(\theta)$ prepares a quantum state $|x(\theta)\rangle$ with $\mathcal{A}\,|x(\theta)\rangle \propto |b\rangle$. The circuit $V(\theta)$ is trained semi-classically via Stochastic Gradient Descend (SGD) to minimize the global cost $C = 1 - |\langle v|\varphi(\theta)\rangle|^2$.

Preliminary experiment

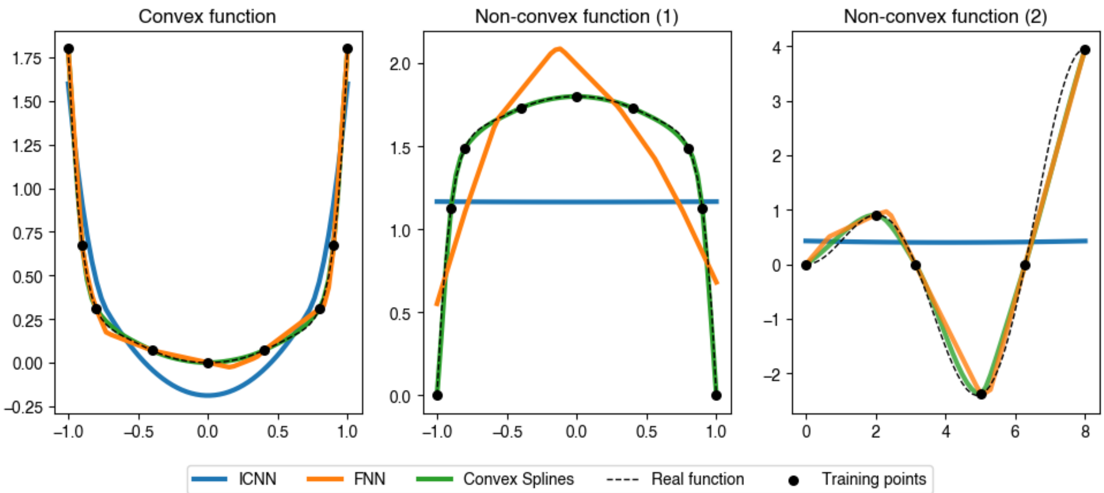

Figure 1: Comparison of ICNN, FNN, linear splines, and convex splines for the interpolation task of three different functions, including a (strictly) convex one (left), a nowhere convex one (center) and one with mixed convexity domains (right). We observe that our method interpolates the curves in the three scenarios, even when the function is non-convex thanks to the SWAM algorithm, while preserving global convexity. The linear interpolation does not preserve convexity due to the lack of added constraints. The FNN approximates the curves without interpolation or convexity preservation. Finally, despite the fact that the ICNN guarantees convexity it is prone to return constant lines and does not interpolate the points. Among the four methods, our algorithm is the only one that interpolates the points while preserving convexity with the added benefits of its mathematical simplicity and easiness of the quantum implementation.

We apply our method for three different types of functions. A strictly convex function $f(x) = 0.4x^2 + 0.5x^4 - 1.6x^6 + 2.5x^{10},\ x \in [-1, 1]$, a nowhere convex function $f(x) = -(0.4x^2 + 0.5x^4 - 1.6x^6 + 2.5x^{10}) + 1.8,\ x \in [-1, 1]$, and a function with mixed behaviour $f(x) = \left(\frac{x}{2}\right)\ \sin(x)\,,\ x \in [0, 8]$. We compare our convex interpolation method against linear interpolation (LI), a Feedforward Neural Network (FNN), and an Input Convex Neural Network (ICNN). The results in Figure 1 show that our convex splines method is the only one that succeeds in interpolating the points while preserving the convexity properties of the subdomains. The generalization of our method from interpolation to regression is done by giving an extra degree of freedom to the set of images S, and tuning them to minimize the global MSE of f when evaluated on the dataset $\mathcal{D}$. Note that the VQLS requires an efficient preparation of the quantum states encoding $\mathcal{A}$, and b which could be challenging in many scenarios. However, the recent advances in quantum state preparation [10, 22, 31] and the constant progress in the area will soon enable the efficient preparation of arbitrary quantum states, making it feasible to apply our method even in those challenging settings.

Acknowledgements

The authors thank Francesco Tacchino, Giulia Cavagnari and Marianna Rapsomaniki for their valuable discussions and feedback regarding the development of the project.

References

[1] Amira Abbas, David Sutter, Christa Zoufal, Aurélien Lucchi, Alessio Figalli, and Stefan Woerner. The power of quantum neural networks. *Nature Computational Science*, 1(6):403–409, 2021.

[2] Sydney N Afriat. The construction of utility functions from expenditure data. *International economic review*, 8(1):67–77, 1967.

[3] Brandon Amos, Lei Xu, and J Zico Kolter. Input convex neural networks. In *International Conference on Machine Learning*, pages 146–155. PMLR, 2017.

[4] Richard Bellman. A markovian decision process. *Journal of mathematics and mechanics*, pages 679–684, 1957.

[5] Stephen P Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[6] Carlos Bravo-Prieto, Ryan LaRose, Marco Cerezo, Yigit Subasi, Lukasz Cincio, and Patrick J Coles. Variational quantum linear solver. *Quantum*, 7:1188, 2023.

[7] Sergey Bravyi, Andrew W Cross, Jay M Gambetta, Dmitri Maslov, Patrick Rall, and Theodore J Yoder. High-threshold and low-overhead fault-tolerant quantum memory. *Nature*, 627(8005):778–782, 2024.

[8] Yann Brenier. Polar factorization and monotone rearrangement of vector-valued functions. *Communications on pure and applied mathematics*, 44(4):375–417, 1991.

[9] Xin-Qiang Cai, Pushi Zhang, Li Zhao, Jiang Bian, Masashi Sugiyama, and Ashley Llorens. Distributional pareto-optimal multi-objective reinforcement learning. *Advances in Neural Information Processing Systems*, 36, 2024.

[10] Almudena Carrera Vazquez and Stefan Woerner. Efficient state preparation for quantum amplitude estimation. *Physical Review Applied*, 15(3):034027, 2021.

[11] Will Dabney, Georg Ostrovski, David Silver, and Rémi Munos. Implicit quantile networks for distributional reinforcement learning. In *International conference on machine learning*, pages 1096–1105. PMLR, 2018.

[12] Will Dabney, Mark Rowland, Marc Bellemare, and Rémi Munos. Distributional reinforcement learning with quantile regression. In *Proceedings of the AAAI conference on artificial intelligence*, volume 32, 2018.

[13] Vojtěch Havlíček, Antonio D Córcoles, Kristan Temme, Aram W Harrow, Abhinav Kandala, Jerry M Chow, and Jay M Gambetta. Supervised learning with quantum-enhanced feature spaces. *Nature*, 567(7747):209–212, 2019.

[14] Matteo Antonio Inajetovic, Filippo Orazi, Antonio Macaluso, Stefano Lodi, and Claudio Sartori. Enabling non-linear quantum operations through variational quantum splines. In *International Conference on Computational Science*, pages 177–192. Springer, 2023.

[15] Leonid V Kantorovich. On the translocation of masses. In *Dokl. Akad. Nauk. USSR (NS)*, volume 37, pages 199–201, 1942.

[16] Youngseok Kim, Andrew Eddins, Sajant Anand, Ken Xuan Wei, Ewout Van Den Berg, Sami Rosenblatt, Hasan Nayfeh, Yantao Wu, Michael Zaletel, Kristan Temme, et al. Evidence for the utility of quantum computing before fault tolerance. *Nature*, 618(7965):500–505, 2023.

[17] Antonio Macaluso, Luca Clissa, Stefano Lodi, and Claudio Sartori. Quantum splines for non-linear approximations. In *Proceedings of the 17th ACM International Conference on Computing Frontiers*, pages 249–252, 2020.

[18] Antonio Macaluso, Filippo Orazi, Matthias Klusch, Stefano Lodi, and Claudio Sartori. A variational algorithm for quantum single layer perceptron. In *International Conference on Machine Learning, Optimization, and Data Science*, pages 341–356. Springer, 2022.

[19] Ashok Makkuva, Amirhossein Taghvaei, Sewoong Oh, and Jason Lee. Optimal transport mapping via input convex neural networks. In *International Conference on Machine Learning*, pages 6672–6681. PMLR, 2020.

[20] Nicola Mariella, Albert Akhriev, Francesco Tacchino, Christa Zoufal, Juan Carlos Gonzalez-Espitia, Benedek Harsanyi, Eugene Koskin, Ivano Tavernelli, Stefan Woerner, Marianna Rapsomaniki, Sergiy Zhuk, and Jannis Born. Quantum theory and application of contextual optimal transport. In *Forty-first International Conference on Machine Learning*, 2024.

[21] Achille Messac and Arun Sivanandan. A new family of convex splines for data interpolation. *Computer aided geometric design*, 15(1):39–59, 1997.

[22] Martin Plesch and Časlav Brukner. Quantum-state preparation with universal gate decompositions. *Physical Review A*, 83(3):032302, 2011.

[23] Blaine Rister and Daniel L Rubin. Piecewise convexity of artificial neural networks. *Neural Networks*, 94:34–45, 2017.

[24] R Tyrrell Rockafellar, Stanislav Uryasev, et al. Optimization of conditional value-at-risk. *Journal of risk*, 2:21–42, 2000.

[25] Ralph Tyrell Rockafellar. *Convex Analysis*. Princeton University Press, Princeton, 1970.

[26] Maria Schuld, Ilya Sinayskiy, and Francesco Petruccione. Prediction by linear regression on a quantum computer. *Physical Review A*, 94(2):022342, 2016.

[27] Ali Siahkamari, Durmus Alp Emre Acar, Christopher Liao, Kelly L Geyer, Venkatesh Saligrama, and Brian Kulis. Faster algorithms for learning convex functions. In *International Conference on Machine Learning*, pages 20176–20194. PMLR, 2022.

[28] Ali Siahkamari, Aditya Gangrade, Brian Kulis, and Venkatesh Saligrama. Piecewise linear regression via a difference of convex functions. In *International conference on machine learning*, pages 8895–8904. PMLR, 2020.

[29] Guoming Wang. Quantum algorithm for linear regression. *Physical review A*, 96(1):012335, 2017.

[30] Ming Yu, Zhuoran Yang, Mladen Kolar, and Zhaoran Wang. Convergent policy optimization for safe reinforcement learning. *Advances in Neural Information Processing Systems*, 32, 2019.

[31] Xiao-Ming Zhang, Tongyang Li, and Xiao Yuan. Quantum state preparation with optimal circuit depth: Implementations and applications. *Physical Review Letters*, 129(23):230504, 2022.