블록 인코딩의 구성: 대안적 QRAM을 중심으로

(Construction of Block-encodings: concerning alternatives of QRAM)

노 현 민 † 김 태 원 ††

(Hyeonmin Roh) (Taewon Kim)

요 약. 블록 인코딩(Block-encoding)은 양자 기계학습 및 선형대수 알고리즘이 제 고전 판본에 대해 보이는 지수적인 속도 증진의 근간인 양자 자료구조, 이른바 QRAM을 구성하는 도구다. 한편 일부 양자 기계학습 및 선형대수 알고리즘들을 오직 다항 시간 감속만으로 고전 알고리즘에 의해 모사할 수도 있는데, 이런 모사를 역양자화(Dequantization)라고 부른다. 또한 역양자화란 QRAM을 구성하는 블록 인코딩의 대수적인 성질에 주목한 결과이기도 하다. 따라서 본고는 역양자화를 방지하려면 블록 인코딩을 어떻게 구성할 것이냐는 문제를 중심으로 블록 인코딩과 인접 개념을 일별한다.

키워드: 양자 기계학습, 양자 알고리즘, 블록 인코딩, 양자 특잇값 변환

Abstract. *Block-encoding* is a tool that constructs the quantum data structure known as QRAM, which serves as the foundation for the exponential speedup observed in quantum machine learning and linear algebra algorithms compared to their classical counterparts. On the other hand, certain quantum machine learning and linear algebra algorithms can be emulated classicaly, in other words, dequantized, with only polynomial-time slowdown. Dequantization is, in part, a result of exploiting the algebraic properties of block-encoding that make up QRAM. Therefore, this paper sketches the notion of block encoding and related concepts, with a central focus on the problem of how to construct it to prevent dequantization.

Keywords: quantum machine learning, quantum algorithms, block-encoding, QSVT

†국립부경대학교 과학컴퓨팅학과 학생

††국립부경대학교 컴퓨터공학과 학생

1 개요

1.1 역사

양자 기계학습(Quantum Machine Learning) 알고리즘은 고전 기계학습 알고리즘에 대해 지수적인 가속(exponential speedup)을 보인다는 주장이 존재한다. 이들 주장과 더불어 양자 기계학습이라는 분야 자체가 연립일차방정식 $Ax = b$의 해를 구하는 HHL 알고리즘[5]에 강하게 의존하는데, 일찍이 Aaronson[1]은 HHL 알고리즘과 양자 기계학습 및 그 이점이 지니는 한계를 지적했다. 머지 않아 Tang[9]은 기계학습에서 주로 사용하는 고전 데이터에 대해 오직 다항 수준의 감속(polynomial slowdown)으로 양자 추천 알고리즘[6]을 모사하는 고전 알고리즘을 내놓으며 이런 과정을 역양자화(dequantization)라고 명명했다.

⋮

1.2 표기 및 기초

2 블록 인코딩

2.1 기초 정의

블록 인코딩은 유니타리가 해밀토니언 시뮬레이션의 효과적인 해법에 관한 Childs et.al [3]의 연구에서 비롯하며, 기본적으로는 유니타리가 아닌 행렬을 양자컴퓨터에서 사용할 수 있도록 하는 기법이다.

행렬 $A \in \mathbb{C}^{r\times c}$가 있다고 하자. 이때 A는 정방일 필요가 없다. 그렇다면 유니타리 U가 존재하여 A가 U의 좌측 상단을 차지한다. 즉 임의의 행렬 성분 $\cdot$ 에 대해 A의 블록 인코딩 U는 아래와 같다.

$$U = \begin{bmatrix} A & \cdot \\ \cdot & \cdot \end{bmatrix} \tag{1}$$

여기서 A의 지수 r, c를 2의 거듭제곱으로 둬, 아래처럼 $|0\rangle$과 $\langle 0|$에 대해 나타내는 것이 목표다.

$$A = (|0\rangle^{\otimes a_L} \otimes I)U(|0\rangle^{\otimes a_R} \otimes I) \tag{2}$$

r, c가 2의 거듭제곱이어야 하는 것은 큰 제약이 아니다. A에 대해 $A_e \in \mathbb{C}^{2^s} \times 2^s$를 정의하여 A를 A_e의 좌측 상단에 두고 나머지 성분은 0으로 두면 되기 때문이다.

일반적으로 블록 인코딩은 [4, 8] 아래처럼 정의한다.

정의 2.1 (블록 인코딩)**.** *A, 조정자 $\alpha \in \mathbb{R}_+$, 블록을 위한 패딩 $a \in \mathbb{R}_+$, 정밀도 $\epsilon \in \mathbb{R}_+$가 주어질 때, U가 아래를 만족하며*

$$\|A/\alpha - (\langle 0|^{\otimes a} \otimes I)U(|0\rangle^{\otimes a} \otimes I)\| \leq \epsilon \tag{3}$$

Q개의 게이트로 구현할 수 있다면 ϵ 정밀도의 α 조정 Q 블록 인코딩이라고 부른다.

의미를 더 직관적으로 소화하기 위해 ϵ과 α를 생략하여 정확하며 1로 조정된 Q 블록 인코딩을 정의하면 아래와 같다.

정의 2.2. *$A \in \mathbb{C}^{r\times c}$가 주어질 때, $U \in \mathbb{C}^{d\times d}$는 U가 $\mathcal{O}(Q)$ 게이트로 구현가능하고 항등행렬의 첫 r열과 c열인 $B_L \in \mathbb{C}^{d\times r}, B_R \in \mathbb{C}^{d\times c}$에 대해*

$$B_L^\dagger U B_R = A \tag{4}$$

를 만족하면 A의 Q 블록 인코딩이라고 부른다.

아래는 이들 정의가 동치라는 사실을 이해하기 위한 예시다.

예시 2.1. $c = 2, d = 8, r = 4$라고 가정하자. 다시 말해 $A \in \mathbb{C}^{4\times 2}, U \in \mathbb{C}^{8\times 8}, B_L \in \mathbb{C}^{8\times 4}, B_R \in$

$\mathbb{C}^{8\times 2}$라고 가정하자. 그렇다면 $B_L^\dagger$는 다음과 같고

$$B_L^\dagger = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

B_R은 아래와 같으며

$$B_R = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

U는 다음과 같다.

$$U = \begin{pmatrix} A & \cdot \\ \cdot & \cdot \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{21} & a_{22} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{31} & a_{32} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{41} & a_{42} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix}$$

이에 아래가 성립한다.

$$\Rightarrow \quad B_L^\dagger U = \begin{pmatrix} a_{11} & a_{12} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{21} & a_{22} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{31} & a_{32} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{41} & a_{42} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix}$$

$$\Rightarrow \quad B_L^\dagger U B_R = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \\ a_{41} & a_{42} \end{pmatrix} = A$$

2.2 QSVT

3 역양자화

3.1 양자 기계학습과 고전 데이터

3.2 양자 선형대수와 고전 선형대수

4 블록 인코딩의 구성

따라서 일부 양자 기계학습 알고리즘은 실질적인 속도 증진을 보이지 않는다. 잇따르는 물음들 가운데 하나가 블록 인코딩에 대한 것이다.

문제 4.1. 역양자화를 방지할 수 있도록 블록 인코딩을 구성하는 다른 방법이 존재하는가*?*

요컨대 Kerenidis와 Praksash[7] 그리고 Chakraborty et.al[2]은 입력 행렬을 저장할 수 있는 크기 혹은 노름을 조정하는 QRAM을 대안으로 내놓는다. 이들 대안은 역양자화될 가능성이 적은데, 왜냐하면 이들 대안이 BQP 완전으로 알려진 [5] 희소 입력 행렬의 모델을 일반화하고 강화하기 때문이다. 이에 대안적인 QRAM상 블록 인코딩이 어떻게 구성되었는지 살핀다.

⋮

4.1 양자적으로 접근 가능한 자료구조

Chakraborty et.al이 내놓은 대안적인 QRAM은 이른바 양자적으로 접근 가능한 자료구조(quantum-accesible data structure)이다. QROM(quantum-random-access read-only memory)에 저장되어 있을 경우, 중첩으로 접근 가능하지만 양자 상태로 저장될 필요는 없는 사실상 고전적인 자료구조다.

⋮

정리 4.1. $A \in \mathbb{C}^{m\times n}$이라고 두자.

(i) $p \in [0,1]$을 고정한다. $A \in \mathbb{C}^{m\times n}$, $A^(p)$, $A^{(1-p)^\dagger}$가 모두 양자적으로 접근 가능한 자료구조상에 저장되어 있을 경우, $\mathcal{O}(polylog(mn/\epsilon))$ 시간 안에 구현할 수 있는 유니타리 U_R과 U_L이 존재하여 $U_R^\dagger U_L$이 A의 $(\eta_{p(A)}, \lceil\log(n+m+1)\rceil, \epsilon)$ 블록 인코딩으로 성립한다.

(ii) 다른 한편 A가 양자적으로 접근 가능한 자료구조상에 저장되어 있을 경우, $\mathcal{O}(\|A\|_F, \lceil\log(m+n)\rceil, \epsilon)$ 시간 안에 구현할 수 있는 유니타리 U_R, U_L이 존재하여 A의 $(\|A\|_F, \lceil\log(m+n)\rceil, \epsilon)$ 블록 인코딩으로 성립한다.

References

[1] Scott Aaronson. Read the fine print. *Nature Physics*, 11:291 – 293, 2015.

[2] Shantanav Chakraborty, András Gilyén, and Stacey Jeffery. The power of block-encoded matrix powers: improved regression techniques via faster hamiltonian simulation. In *International Colloquium on Automata, Languages and Programming*, 2018.

[3] Andrew M. Childs, Robin Kothari, and Rolando D. Somma. Quantum algorithm for systems of linear equations with exponentially improved dependence on precision. *SIAM J. Comput.*, 46(6):1920–1950, jan 2017.

[4] András Gilyén, Yuan Su, Guang Hao Low, and Nathan Wiebe. Quantum singular value transformation and beyond: exponential improvements for quantum matrix arithmetics. pages 193–204, 06 2019.

[5] Aram W. Harrow, Avinatan Hassidim, and Seth Lloyd. Quantum algorithm for linear systems of equations. *Phys. Rev. Lett.*, 103:150502, Oct 2009.

[6] Iordanis Kerenidis and Anupam Prakash. Quantum Recommendation Systems. In Christos H. Papadimitriou, editor, *8th Innovations in Theoretical Computer Science Conference (ITCS 2017)*, volume 67 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 49:1–49:21, Dagstuhl, Germany, 2017. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

[7] Iordanis Kerenidis and Anupam Prakash. Quantum gradient descent for linear systems and least squares. *Phys. Rev. A*, 101:022316, Feb 2020.

[8] Patrick Rall. Quantum algorithms for estimating physical quantities using block encodings. *Phys. Rev. A*, 102:022408, Aug 2020.

[9] Ewin Tang. A quantum-inspired classical algorithm for recommendation systems. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2019, page 217–228, New York, NY, USA, 2019. Association for Computing Machinery.