

Quantum algorithm for linear system of equations (Harrow, Hassidim, Lloyd, 2008) 해설

계산과학 2: 노현민-김태원 조

2023년 9월 8일

$N \times N$ 에르미트(Hermitian) 행렬 A 가 있다고 하자. 에르미트 행렬이란 켤레전치(conjugate transpose)가 자기 자신과 같은 복소정방행렬(complex square matrix)이다. 또한 단위벡터(unit vector) \mathbf{b} 가 있다고 하자. 단위벡터란 길이 혹은 노름(norm)이 1인 벡터다. 이에 연립일차 방정식을 \mathbf{x} 에 대해 푼다. 다시 말해 아래 식에서 \mathbf{x} 를 구한다.

$$A\mathbf{x} = \mathbf{b}$$

이를테면 아래 연립일차방정식에서 \mathbf{x} 는 (b_1, b_2, b_3) 이다. 보통 기초선형대수학 수업에서는 대각화와 가우스-조르단 소거법을 사용해 푼다.

$$\begin{aligned} \begin{aligned} 1x_1 + 0x_2 + 0x_3 &= b_1 \\ 0x_1 + 1x_2 + 0x_3 &= b_2 \\ 0x_1 + 0x_2 + 1x_3 &= b_3 \end{aligned} &\longrightarrow \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} x_1 + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} x_2 + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} x_3 = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \\ &\longrightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \\ &\longrightarrow A\mathbf{x} = \mathbf{b} \end{aligned}$$

\mathbf{b} 를 아래와 같은 양자상태로 나타낸다.

$$|\mathbf{b}\rangle = \sum_{i=1}^N b_i |i\rangle$$

앞선 예시로 켓(ket) 표기 $|i\rangle$ 을 설명하면 아래와 같다.

$$|\mathbf{b}\rangle = \sum_{i=1}^{N=3} b_i |i\rangle = b_1 |1\rangle + b_2 |2\rangle + b_3 |3\rangle = b_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + b_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + b_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

여기에 **해밀토니언 시뮬레이션** Hamiltonian simulation이라는 기법을 적용한다.

해밀토니언 시뮬레이션

해밀토니언 시뮬레이션이란 양자정보학의 기초 문제 가운데 하나로, 양자계 quantum systems의 시뮬레이션에 요구되는 계산복잡도와 알고리즘을 묻는다. 다시 말해 양자 계의 (시간의 흐름을 따른) 진화 혹은 변화 evolution를 구현하는 알고리즘을 구하는 문제다. 유명한 파인만 Richard Feynman 씨가 1982년에 고안했으며, 해밀토니언의 고전적 시뮬레이션에는 계의 상태에 따른 지수적인 exponential 감속이 있기 때문에 아무래도 양자컴퓨터 같은 게 필요하지 않겠냐는 맥락이었다.

여기서 해밀토니언 H 란 수학적으로 말하면 그냥 큐비트의 개수 n 에 대한 $2^n \times 2^n$ 에르미트 행렬이고, 물리학적으로 말하면 어느 계의 총 에너지에 대응하는 연산자다. 시간 t 에 따른 해밀토니언 H 의 이상적인 변화를 e^{-iHt} 로 나타낸다.

오일러 공식 $e^{i\theta} = \cos \theta + i \sin \theta$

복소평면, 다시 말해 x 좌표가 복소수 z 의 실수부분($\text{Re}(z)$)을 나타내고 y 좌표가 z 의 허수부분($\text{Im}(z)$)인 좌표평면상에서 모든 복소수 z 를 아래처럼 나타낼 수 있다.

$$z = re^{i\theta} = r(\cos \theta + i \sin \theta)$$

대충 말하면 복소평면상의 한 단위를 θ 만큼 돌리고 r 만큼 늘리면 모든 z 를 나타낼 수 있고 이런 z 가 평면상에서는 $\cos \theta + i \sin \theta$ 를 r 만큼 늘린 꼴이다. ($\cos \theta + i \sin \theta$ 가 평면상의 한 점을 나타낸다는 사실은 그림으로 그리면 효과적으로 이해할 수 있고 아무튼) 이를 복소수의 **극 형식** polar form이라고 한다. 또한 여기서 단위 복소수 $e^{i\theta}$ 는 따로 **위상 인자** phase factor라고 부른다.

위상 인자

양자역학에서 위상 인자란 $|\psi\rangle$ 와 $|\phi\rangle$ 에 곱하는 복소 계수 $e^{i\theta}$ 다. 단위 복소수이기 때문에 곱하더라도 전역 위상 global phase에서는 아무 일이 일어나지 않는다. 다시 말해 아래가 성립한다.

$$e^{i\theta}(|0\rangle + |1\rangle) = |0\rangle + |1\rangle$$

하지만 상대 위상 relative phase $e^{i(\theta-\theta')}$ 가 있을 수 있으므로 아래 또한 성립한다.

$$e^{i\theta}|0\rangle + e^{i\theta'}|1\rangle \neq |0\rangle + |1\rangle$$

이 e^{-iHt} 와 최대 오류 ϵ 에 대해 $\|U - e^{-iHt}\| \leq \epsilon$ 을 만족하는 유니타리 변환 U 에

근사하는 알고리즘을 찾는 문제가 바로 해밀토니안 시뮬레이션이다.

유니타리

복소정방행렬 U 는 켈레진치 U^\dagger 와 곱할 때 항등행렬 I 가 나오면 유니타리라고 한다. 다시 말해 $U^\dagger U = U U^\dagger = I$ 를 만족하는 U 는 유니타리다.

다시 말해 $|b\rangle$ 를 시간 t 에 따른 (일종의) 해밀토니안 A 의 이상적인 변화 e^{iAt} 로 다룬다. (아무튼 A 라는 계에 x 라는 변환을 취한 결과가 b 이기 때문이고) 여기서 핵심은 A 를 지수로 다뤘다는 사실이다. 이 사실은 **위상 추정** phase estimation이라는 기법을 통해 $|b\rangle$ 를 분해할 수 있도록 한다.

위상 추정 알고리즘

주어진 유니타리 연산자의 고윳값에 대응하여 위상을 추정하는 알고리즘이다.

고윳값

체 혹은 스칼라 집합 K 에 대해 벡터공간 V 상의 선형변환 $T: V \rightarrow V$ 가 주어질 때, $v \in V$ 와 $\lambda \in K$ 가 존재하여

$$v \neq 0 \quad \& \quad Tv = \lambda v$$

를 만족하면 v 는 T 의 고유벡터이며 λ 는 T 의 고윳값이다.

또한 유니타리 행렬에 대한 스펙트럼 정리에 (<https://www.math.purdue.edu/~eremenko/dvi/lect3.26.pdf>) 따르면, 유니타리 행렬의 모든 고윳값은 절댓값 1, 다시 말해 단위율unit modulus을 지닌다. 따라서 고윳값은 위상으로 특징지어지고, 알고리즘이 고윳값을 통해 위상을 추정할 수 있다.

A 의 고유기저 u_j 와 그 고윳값 λ_j 그리고 $|b\rangle = \sum_{j=1}^N \beta_j |u_j\rangle$ 에 대해 Ax 를 아래처럼 나타낼 수 있다.

$$\sum_{j=1}^N \beta_j |u_j\rangle |\lambda_j\rangle$$

그리고 정규화 상수 C 에 대한 선형사상 $f: |\lambda_j\rangle \mapsto C\lambda_j^{-1} |\lambda_j\rangle$ 를 적용하면 아래처럼 $|x\rangle$ 를 구할 수 있다.

$$\sum_{j=1}^N \beta \lambda_j^{-1} |u_j\rangle = A^{-1} |b\rangle = |x\rangle.$$

물론 여기서 f 는 선형사상일 뿐 유니타리 변환이 아니다. 따라서 실패할 수도 있다. 또한 HHL 알고리즘의 성능을 결정하는 요인은 A 의 최대 고윳값과 최소 고윳값 간의 비율을 나타내는 A 의 조건수condition number k 다. k 가 증가하면 A 에 역행렬이 존재하지

않을 확률도 증가한다. 이에 HHL 알고리즘은 A 에서 역행렬이 존재하는 이른바 특이행렬(singular matrix)들의 값이 $1/k$ 와 1 사이의 값이라고 가정한다. 다시 말해 다음을 가정한다.

$$\frac{1}{k^2} \leq A^\dagger A \leq I$$

그리하여 HHL 알고리즘의 실행시간은 $|x\rangle$ 상태의 출력 과정상 취할 수 있는 오류 ϵ 에 대해 $k^2 \log(N)/\epsilon$ 에 비례한다.

여기까지 HHL 알고리즘이 취하는 로그시간 해법을 개괄했다. 애론슨(Scott Aaronson)은 몇 가지 결정적인 가정을 지적한다.

첫째, \mathbf{b} 라는 벡터를 양자상태 $|b\rangle = \sum_{i=1}^N b_i |i\rangle$ 로 어떻게든 재빠르게 읽어들이어야 한다. 벡터상의 성분들을 양자중첩으로 한 번에 읽는 “양자 RAM” 혹은 QRAM 같은 장치를 가정해서 어떻게든 재빠르게 읽어들이지 않으면 결국 $|b\rangle$ 를 준비하는 작업에만 어떤 상수 c 에 대해 N^c 스텝이 걸린다. 이미 로그시간 해법에서 벗어났다.

둘째, 마찬가지로 $|b\rangle$ 를 시간 t 에 따른 해밀토니안 A 의 이상적인 변화 e^{iAt} 로 다루는 작업에 N^c 스텝이 걸리지 않으려면 각 i 에 대해 QRAM이 존재하여 재빠르게 저장해야 하며, A 가 희소(sparse) 행렬, 다시 말해 A 가 어떤 고정된 정수 s 에 대해 행마다 0이 아닌 성분은 s 개만 지녀야 한다.

셋째, A 는 역행렬의 존재에 의해 그저 가역행렬이기만 해서는 안 된다. 말하자면, 강력하게 가역이어야 한다. 즉 앞서 언급한 $k = |\lambda_{\max}/\lambda_{\min}|$ 가 N^c 꼴을 보이면 말짱 도루묵이다.

넷째, 해 벡터 $x = (x_1, \dots, x_n)$ 을 나타내는 작업에만 n 스텝이 요구된다. HHL 알고리즘은 x 자체가 아니라 $\log n$ 큐비트의 양자상태 $|x\rangle$ 를 출력하는데, 특정 성분 x_i 를 측정하려면 알고리즘을 대략 n 번 반복해야 하기에, 지수적인 가속이 실패한다.

요약하면, 이들 조건을 전부 지키면서 HHL 자체를 응용할 구석은 많지 않다. 다만 HHL은 여타 양자기계학습 알고리즘에 대해 템플릿 역할을 맡을 수 있다. $|b\rangle$ 를 어떻게 마련할 것인가? e^{-iAt} 를 어떻게 적용할 것인가? $|x\rangle$ 를 어떻게 측정할 것인가? 이들 물음에 답하는 것이 HHL 이후 양자기계학습 알고리즘들의 과업이다.

HHL의 요는 실상 $Ax = b$ 의 해를 구하는 로그시간 방법이 아니라 임의의 양자알고리즘을 2^n 개의 선형방정식으로 이루어진 연립일차방정식으로 시뮬레이션하여 속도 증진을 이룰 수 있다는 것이다. 이에 탕(Ewin Tang)이 반문하기를, 이런 시뮬레이션을 고전 컴퓨터가 시뮬레이션할 수 있다면, 여기서 “양자”라는 표현에는 무슨 의미가 있는가?