계산과학2 (1)

노현민-김태원 조

2023년 9월 19일

차 례

제 1 장

HHL

HHL은 양자기계학습이라는 분야 일반과 더불어 양자기계학습이 고전기계학습에 대해 속도 증진을 취한다는 주장의 바탕을 이루는 양자 알고리즘으로, 연립일차방정식의 해가 지니는 특성을 추정한다.

고전 알고리즘은 선형방정식 N개의 해를 근사하는 작업에만 N에 비례하는 시간을 요구하고 해를 작성하는 시간에는 N에 따른 시간을 요구한다. 하지만 부분집합의 가중합만이 필요한 기계학습에서 그렇듯이 우리는 연립일차방정식의 전체 해가 아니라 해의 어느 함수를 계산하고자 할 수도 있다. HHL의 요는 양자컴퓨터가 그러한 함수의 값을 N의 로그에 비례하는 시간과 조건수에 대해 다항시간으로 근사할 수 있다는 것이다. 이때 조건수는 시스템이 불량하거나 민감한 정도를 나타내는 척도로, HHL의 경우 N에는 지수적인 속도 증진을 취하며 조건수에는 비슷하고 오류에는 고전 컴퓨터보다 더 나쁜 성능을 보인다. 따라서 HHL은 N이 크고 조건수가 작다면 충분히 유용한 양자 알고리즘이라는 것이 Harrold, Hassidim, Lloyd의 주장이다.

에르미트 $N \times N$ 행렬 A, 단위 벡터 $\mathbf{b}$가 있다고 하자. 이때 에르미트 행렬이란 켤레전치가 자기 자신과 같은 복소정방행렬을 말한다. 연립일차방정식의 해를 구하는 문제는 아래와 같은 $\mathbf{x}$를 찾는 것이다.

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

우선 $\mathbf{b}$를 아래와 같은 양자 상태로 나타낸다.

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

이에 해밀토니언 시뮬레이션이라는 기법을 사용해 e^{iAt}를 상이한 시간 t의 중첩 $|b\rangle$에 적용한다. e^{-iHt}는 계의 총 에너지에 대응하는 연산자인 해밀토니언 H의 시간 t에 따른 이상적인 변화를 말하는데, 해밀토니언 시뮬레이션은 최대 오류 ϵ에 대해 $\|U - e^{-iHt}\| \leq \epsilon$

인 유니타리 변환 U에 근사하는 알고리즘을 찾는 문제다. 따라서 행렬 A를 해밀토니언으로 두고 시간 t에 따른 복소평면상의 매끄럽고 이상적인 변화 결과 e^{-iAt}를 $|b\rangle$에 적용하는 것, 다시 말해 $|b\rangle$를 A의 고유벡터 $u_j = e^{-iAt}$로 분해하여 이에 대응하는 고윳값 λ_j를 취하는 것이다. 이후 $|b\rangle = \sum_{j=1}^{N} \beta_j |u_j\rangle$에 대해 연립일차방정식은 아래와 같은 꼴을 보인다.

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

이어서 정규화 상수 C에 대해 $|\lambda_j\rangle$를 $C\lambda^{-1} |\lambda_j\rangle$로 보내는 선형사상을 적용한다. 물론 이때 이 연산이 유니타리가 아니기 때문에 이 과정에서 실패가 발생할 수 있다. 앞서 e^{-iAt}를 고유벡터 삼아 고윳값 λ_j를 추정하는 기법을 위상 추정이라고 부르는데, 이 기법이 성립하려면 연산자가 유니타리여(서 단위율을 지녀)야 하기 때문이다. 아무튼 성공하면 $|\lambda_j\rangle$ 항목을 uncompute해서 아래에 상응하는 상태를 취한다.

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

이때 유니타리 행렬의 모든 고윳값은 절댓값 1, 단위율을 지니기 때문에 유니타리 연산자의 고윳값에 위상을 대응시킬 수 있다. 그리하여 A의 고유기저 u_j와 고윳값 λ_j 그리고 $|b\rangle = \sum_{j=1}^{N} \beta_j |u_j\rangle$에 대해 Ax를 아래처럼 나타낼 수 있다.

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

$N \times N$ 에르미트Hermitian 행렬 A와 벡터 $\mathbf{b}$가 있다. 여기서 에르미트 행렬이란 켤레전치conjugate transpose가 자기 자신과 같은 복소정방행렬complex square matrix을 뜻한다. 이에 연립일차방정식을 $\mathbf{x}$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 수업에

서는 대각화와 가우스-조르단 소거법을 사용해 푼다.

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

1.1 양자상태 표현과 브라-켓 표기

b를 아래와 같은 양자상태로 나타낸다.

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

앞선 예시로 켓ket 표기 '$|\cdot\rangle$'를 도입하면 아래와 같다.

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

1.2 해밀토니언 시뮬레이션: 오일러 공식, 위상 인자, 유니타리

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

해밀토니언 시뮬레이션

해밀토니언 시뮬레이션이란 양자정보학의 기초 문제 가운데 하나로, 양자계quantum systems의 시뮬레이션에 요구되는 계산복잡도와 알고리즘을 묻는다. 다시 말해 양자계의 (시간의 흐름에 따른) 진화 혹은 변화evolution를 구현하는 알고리즘을 구하는 문제다. 파인만Richard Feynman이 해밀토니안의 고전적 시뮬레이션은 계의 상태에 따른 지수적인exponential 감속을 요구하니 아무래도 양자컴퓨터 같은 게 필요하지 않겠냐며 1982년 고안한 문제다.

여기서 해밀토니안 H란 수학적으로 말하면 큐비트의 개수 n에 대한 $2^n \times 2^n$ 에르미트 행렬이고, 물리학적으로 말하면 계의 총 에너지에 대응하는 연산자다. 시간 t에 따른 해밀토니안 H의 이상적인 변화를 e^{-iHt}로 나타낸다.

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

복소평면, 다시 말해 x좌표가 복소수 z의 실수부분(Re(z))을 나타내고 y좌표가 z의 허수부분(Im(z))인 좌표평면상에서 모든 복소수 z를 아래처럼 나타낼 수 있다.

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

대충 말해 복소평면상의 한 단위를 θ만큼 돌리고 r만큼 늘리면 모든 z를 나타낼 수 있고 이런 z가 평면상에서는 $\cos\theta + i\sin\theta$를 r만큼 늘린 꼴이라는 뜻이다. 이를 복소수의 **극 형식polar form**이라고 한다. 또 여기서 단위 복소수 $e^{i\theta}$는 따로 **위상 인자phase factor**라고 부른다.

위상 인자

양자역학에서 위상 인자란 $|\psi\rangle$와 $\langle\phi|$에 곱하는 복소 계수 $e^{i\theta}$다. 단위 복소수이므로 곱하더라도 *전역 위상global phase*에서는 아무 일이 일어나지 않는다. 다시 말해 $e^{i\theta}(|0\rangle + |1\rangle)$와 $|0\rangle + |1\rangle$이 동치다. 하지만 이른바 *상대 위상relative phase* $e^{i(\theta-\theta')}$이 있을 수 있으므로 $e^{i\theta}|0\rangle + e^{i\theta'}|1\rangle$와 $|0\rangle + |1\rangle$은 동치가 아니다.

이 e^{-iHt}와 최대 오류 ϵ에 대해 $\|U - e^{-iHt}\| \leq \epsilon$인 유니타리 변환 U과 근사하는 알고리즘을 찾는 문제가 바로 해밀토니안 시뮬레이션이다.

유니타리

복소정방행렬 U는 켤레전치 $U^\dagger$와 곱할 때 항등행렬 I가 나오면 유니타리라고 한다. 즉 $U^\dagger U = UU^\dagger = I$를 만족하는 U는 유니타리다.

그리하여 $|b\rangle$를 시간 t에 따른 (일종의) 해밀토니안 A의 이상적인 변화 e^{iAt}로 다룬다. (아무튼 A라는 계에 $\mathbf{x}$라는 변환을 취한 결과가 $\mathbf{b}$이기 때문이고) 여기서 핵심은 A를 지수로 다룰 수 있다는 사실이다.

1.3 위상 추정 알고리즘: 유니타리 행렬의 고윳값

이 사실은 **위상 추정phase estimation**이라는 기법으로 $|b\rangle$를 분해하도록 한다.

위상 추정 알고리즘

주어진 유니타리 연산자의 고윳값에 대응하여 위상을 추정하는 알고리즘이다.

고윳값

체 혹은 스칼라 집합 K에 대해 벡터공간 V상의 선형변환 $T : V \to V$가 주어질 때, $v \in V$와 $\lambda \in K$가 존재하여

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

를 만족하면 v는 T의 고유벡터이며 λ는 T의 고윳값이다.

또한 유니타리 행렬에 대한 스펙트럼 정리에 따르면, (그렇다는 사실을 그냥 받아들이면) 유니타리 행렬의 모든 고윳값은 절댓값 1, 다시 말해 단위율unit modulus을 지닌다. 따라서 고윳값이 위상을 특성으로 지녀 알고리즘은 고윳값으로 위상을 추정할 수 있다.

A의 (고유벡터로 구성한 기저인) 고유기저 u_j와 그 고윳값 λ_j 그리고 $|b\rangle = \sum_{j=1}^{N} \beta_j |u_j\rangle$에 대해 Ax를 아래처럼 나타낼 수 있다.

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

그리고 정규화 상수 C에 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 f는 선형사상일 뿐 유니타리 변환이 아니다. 따라서 실패할 수도 있다.

1.4 조건수 k와 HHL의 로그시간 성능

HHL 알고리즘의 성능을 결정하는 요인은 A의 최대 고윳값과 최소 고윳값 간의 비율을 나타내는 A의 조건수condition number k다. k가 증가하면 A에 역행렬이 존재하지 않을 확률도 증가한다. 이에 HHL 알고리즘은 역행렬이 존재하는 이른바 특이행렬singular matrix들의 값을 1/k와 1 사이의 값으로 가정한다. 즉 다음을 가정한다.

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

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

1.5 HHL의 결정적인 가정들과 "양자"기계학습의 기초

여기까지 HHL 알고리즘이 취하는 로그시간 해법을 개괄했다. 애론슨Scott Aaronson은 몇 가지 결정적인 가정을 지적한다.

첫째, $\mathbf{b}$라는 벡터를 양자상태 $|b\rangle = \sum_{i=1}^{N} b_i |i\rangle$로 어떻게든 재빠르게 읽어들여야 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 c에 대해 N^c 스텝이 걸린다. 이미 로그시간 해법에서 벗어났다. 그리고 QRAM이 있더라도 $\mathbf{b}$는 대체로 균일한 벡터여야 한다. 즉 어떤 b_i가 나머지보다 특별히 커서는 안 된다.

둘째, $|b\rangle$를 시간 t에 따른 해밀토니안 A의 이상적인 변화 e^{iAt}로 다루는 작업은 N^c 스텝이 걸리지 않으려면 각 i에 대해 QRAM이 존재하여 재빠르게 저장해야 하며 A가 희소sparse 행렬, 다시 말해 A가 어떤 고정된 정수 s에 대해 행마다 0이 아닌 성분은 s개만 지녀야 한다.

셋째, A는 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 $x = (x_1, \ldots, x_n)$을 나타내는 작업에만 n 스텝이 요구된다. HHL 알고리즘은 x 자체가 아니라 $\log n$ 큐비트의 양자상태 $|x\rangle$를 출력하는데, 특정 성분 x_i를 측정하려면 알고리즘을 대략 n번 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^n개의 선형방정식으로 이루어진 연립일차방정식으로 *시뮬레이션*하여 속도 증진을 이룰 수 있다는 것이다. 이에 탕Ewin Tang이 반문하기를, 이런 시뮬레이션을 고전컴퓨터가 시뮬레이션할 수 있다면, 여기서 "양자"라는 표현에는 무슨 의미가 있는가?

제 2 장

추천 알고리즘