

量子コンピュータを用いた変分アルゴリズムと機械学習



御手洗光祐

大阪大学大学院基礎工学研究科
mitarai@qc.ee.es.osaka-u.ac.jp

藤井啓祐

大阪大学大学院基礎工学研究科
fujii@qc.ee.es.osaka-u.ac.jp

ファインマンが指摘したように、量子系は粒子数に対して指数的に自由度が増えるため、従来コンピュータでのシミュレーションは難しくなる。量子力学の原理によって計算を行う量子コンピュータは、このような問題を根本的に解決することができるかと期待されている。今では、原子や電子、そして量子化された電気回路など、個々の量子系を精密に制御する量子エレクトロニクス技術の進展に伴って、量子コンピュータの実現が現実味を帯びつつある。現在、Google、IBM、Microsoftなどの巨大IT企業に加え、ベンチャー企業であるRigetti computingなどが量子コンピュータの実現に向けた基礎研究開発を行っており、すでに数十量子ビット程度の規模の量子コンピュータが実現している。さらに、ここ数年で数百量子ビット規模の量子コンピュータが実現されると期待されている。十分に精度の高い操作が可能で100量子ビット程度の系ができれば、スーパーコンピュータを用いても完全にシミュレーションすることは難しく、このような規模でも大きな潜在能力を秘めている。一方で、量子誤り訂正を実装できるほどの規模がないため、複雑な量子アルゴリズムを実装することはできない。このようなデバイスは、Noisy Intermediate Scale Quantum (device) を略しNISQデバイスと呼ばれている。

NISQデバイスはノイズの影響から、多くのステップを要するような複雑な量子計算を実行することができない。できるだけ、量子計算をコンパクトに設計し有効活用する方法が求められており、物理分野でも古くから行われてきた、変分法による量子系の解析が注目を集めている。従来のコン

ピュータ上で実行する変分法では、量子状態を効率のよい試行関数によって表現する必要がある。密度行列繰り込み群やテンソルネットワークなど、計算コストをできるだけ抑えて量子相関を取り込む方法が検討されてきた。それに対して、量子コンピュータ上の量子状態を試行関数として利用すれば、複雑な量子相関を取り込んだとしても、その状態の生成からエネルギーや観測量の評価は、物理法則を用いて効率よく実行することができる。まさに、ファインマンが指摘したように、量子力学で動作する量子コンピュータを用いて変分法の試行関数を表現するのである。このようなアプローチで、量子多体系の基底状態を求め変分法が変分量子固有値法 (Variational Quantum Eigensolver, VQE) である。

試行関数によるモデル化は、物理分野だけでなく、ニューラルネットワークなどの機械学習においてもよく行われている。最近では、量子多体系をニューラルネットワークで表現したり、物質相を機械学習で分類する、といったアプローチの研究も進んでいる。我々の最近の研究では、教師あり機械学習に量子回路を用いた変分的なアプローチを利用することを提案している。量子ビットに対して指数的に大きな次元となるヒルベルト空間を機械学習のための特徴量空間として利用しようという試みである。また、このような量子コンピュータを用いた機械学習を基底状態などの量子系を学習するために応用する研究も行っている。

このように、変分アプローチによる量子古典ハイブリッドアルゴリズムは、NISQデバイスでの利用が期待されており、発展が期待される分野である。

—Keywords—

NISQデバイス：

Noisy intermediate scale quantum デバイスの略。ここ数年で実現されると思われる量子コンピュータ、すなわち数十～数百量子ビットを備えるがノイズの影響を無視できないようなデバイスのこと。～100量子ビットの量子系をシミュレートするのがスーパーコンピュータを用いても難しいことから、ある意味で古典計算機を超えた計算能力を持っていると考えられる。この計算能力を活かすためのアルゴリズムが探索されている。

変分量子アルゴリズム：

量子コンピュータを用いて適当なパラメータを持った波動関数を作り出し、その波動関数に依存する何らかの目的関数を逐次的に最小(最大)化するアルゴリズムのこと。パラメータの最適化ルーチン自体は古典計算機上で走らせ、量子コンピュータは波動関数の生成と物理量の測定に徹する。古典計算機ではそもそも表現できない波動関数を扱える可能性があり、特に量子系の解析において注目を集めている。

1. はじめに

量子コンピュータとは、量子力学の原理に基づいて計算を行うコンピュータであり、その原点は、1980年代にまで遡る。量子電磁気学への貢献でノーベル賞を受賞したファインマンは、“Nature isn't classical, ... if you want to make a simulation of Nature, you'd better make it quantum mechanical”と指摘している。また、それと同時に、“it's a wonderful problem, because it doesn't look so easy”とも言った。80年代当時の感覚では、1つ1つの原子やスピンなどを量子的に操作し測定する、というのは全く現実的ではなかったであろう。しかし21世紀に入り、光子、原子、電子、そして量子化された電気回路など、様々な物理系において、個々の量子系を精密に制御する、量子エレクトロニクス技術が発展してきた。2012年には、真空中に捕獲したイオンを用いて量子状態を制御するイオントラップ技術を開発したワインランド (D. J. Wineland)、そして、空洞共振器内の量子化された電磁波と原子との相互作用による量子状態制御を実現したアロシュ (S. Haroche) らにノーベル賞が授与されている。また、1999年、当時NECにいた中村・蔡らによって世界で初めて実現された超伝導量子ビットは、20年の時を経て、現在量子コンピュータの最有力候補として集積化に向けた研究が進められている。¹⁾ また、2014年以降、GoogleやIBM、そしてベンチャー企業のRigetti computingらが参入してくるによって研究開発が加速されており、現在20量子ビット (空間の次元にして 2^{20} 次元) が高い精度で自在に制御可能になってきている。さらに、50-100量子ビット規模の量子コンピュータの実現にむけた取り組みが行われている。

しかしながら、依然として現実と理想とのギャップは大きい。理論家が夢想する完全な量子コンピュータの実現には、ノイズを克服するための量子誤り訂正機能の搭載が不可欠であり、さらに20年はかかると予想されている。現在実現されている量子コンピュータに量子誤り訂正機能を搭載させると1つの守られた量子ビットを作るのがせいぜいで、量子情報を守りながら計算をすることはできない。一方で、地道な実験家の努力のおかげで、ノイズの原因となる物理の理解は進み、量子ビットの寿命や量子操作の精度はますますよくなってきている。50量子ビットの量子計算は、従来コンピュータで愚直にシミュレーションすると、 2^{50} 個の複素数を確保する必要があり、単純にこの容量を計算すると、これは単精度小数を用いてもペタバイトのオーダーのメモリが必要となる。もう少し洗練されたシミュレーション方法を用いたとしても、十分高い精度の100量子ビットの量子コンピュータを従来コンピュータでシミュレートすることは難しそうだとされている。このように、誤り耐性のある量子コンピュータよりは規模が小さくノイズも含まれるが、従来コンピュータによるシミュレーションが難しくなるようなレベル、というのが現在の量子コンピュータの立ち位置であり、NISQ (Noisy Inter-

mediate Scale Quantum) 時代と呼ばれている。²⁾

NISQ時代の量子コンピュータが抱えるミッションは大きくわけて3つある。1つ目は、量子コンピュータによって計算が加速されるという現象 (量子超越性) を定量的に検証することである。人類は未だ量子力学を用いて計算が加速されるという現象を観測していない。これを情報科学的な緻密な議論に基づいて検証しようという試みである。^{3,4)} 2つ目は、夢の大規模な量子コンピュータの実現にむけた問題点を洗い出すためのテストベッドとしての役割である。50量子ビットが実現されれば量子情報をノイズから守る、量子誤り訂正の実証実験を行うことができる。この実証実験からさらなる大規模化に必要な技術的問題点を洗い出し、解決方法を見出す必要がある。3つ目は、その量子超越性を活かした応用計算だ。応用分野としては、量子力学が有利に働くであろうと予想される、物性系や分子 (化学系) の量子シミュレーションに始まり、機械学習、近似最適化に至るまで多岐にわたっている。従来の量子コンピュータ上のアルゴリズムは、高度に作り込まれたものであり、その結果、量子的な計算の加速についての緻密な議論を展開することができる。その一方、操作が複雑すぎてNISQ時代の量子コンピュータでは動かすことができない。NISQ時代の量子コンピュータは、ノイズの影響によってできるだけ速く計算を終わらせる必要があり、許された量子操作の回数も限定される。このような状況で、動作する量子アルゴリズムが最近盛んに研究され始めた。興味深いことに、このような現実的な問題と折り合いを付けることによって、新たに従来の物理的な手法との接点や類似性が出始めている。

本稿では、このようなNISQ時代の量子コンピュータを物性計算、量子化学計算、そして機械学習等に活用するための量子古典ハイブリッドアルゴリズムについて解説し、他の物理分野の接点について眺めてみたい。

2. 変分量子固有値法 (VQE)

物性物理や量子化学において、基底状態のエネルギーを求めたり物性値を計算することは重要な問題である。厳密対角化や量子モンテカルロ法など様々なアプローチが研究されている。変分法によるアプローチでは、古くから知られているレイリー・リッツ法に始まり、量子多体系の相関をうまく取り込んだ試行関数を用いる密度行列繰り込み群、⁵⁾ そしてテンソルネットワーク法⁶⁾ などの研究が知られている。複雑な量子相関 (エンタングルメント) を取り込むと一般に計算コストは大きくなり、厳密解を得るために必要な計算資源は、対象とする物理系のサイズに対して指数関数的に増えていく。

量子コンピュータ上では、量子状態の表現や観測量の評価などを、効率よく行うことができる。量子ビットの集合とみなせるスピン1/2系の量子状態はもちろんのこと、ボソンやフェルミオンのものについても、適切な変換を行え

ば、考えたいヒルベルト空間の次元 D に対して $\log_2 D$ 程度の量子ビットで表現可能である。変分量子固有値法 (Variational Quantum Eigensolver, VQE)⁷⁻⁹⁾ は、量子コンピュータに対する量子ゲート操作を用いて、量子状態として試行関数を構築する。この試行状態は、一般に古典コンピュータでは表現できないような状態も含む。そして、その状態からエネルギーを測定し、試行関数のパラメータ (量子操作に含まれるパラメータ) をエネルギーを下げる方向に更新することによって、基底状態とそのエネルギーを求める方法である。VQE は、量子コンピュータ実機を用いた簡単な実証実験とともに2014年に提案された比較的新しいアプローチである。測定したエネルギーやその勾配情報から、パラメータの更新方法を決める部分は従来のコンピュータで処理することから、量子古典ハイブリッド量子アルゴリズムの一種とされており、すべて量子コンピュータ上で処理をしエネルギーを求める従来の量子位相推定アルゴリズムに比べると小規模のNISQデバイス上でも動作するという利点がある。

VQE は、与えられたハミルトニアン H の基底状態を変分的に求めるアルゴリズムであり、図1のように古典コンピュータと共同して動作する。 n 量子ビットの量子コンピュータがあったとしよう。物理系としては、 n 個のスピン $1/2$ が張るヒルベルト空間と等価である。量子回路 $U(\theta)$ は、量子ビットの回転の角度などの量子操作を決めるパラメータ θ が含まれており、量子コンピュータ上で効率よく実行できる $2^n \times 2^n$ のユニタリ演算子である。それを初期状態に作用して出力される状態

$$|\psi(\theta)\rangle = U(\theta)|0\rangle^{\otimes n} \quad (1)$$

が変分法の試行関数となる量子状態であり、以降、**試行状態**と呼ぶことにする。

ここでは、ハミルトニアン H がスピン $1/2$ 粒子のパウリ演算子の積で与えられるような場合を考える。対象とする系がフェルミ粒子系であればジョルダン・ウィグナー変換をすればよく、ボーズ粒子系であれば、1つ1つのモードを有限の粒子数状態で打ち切り、1モードを複数の量子ビットを用いて表現すればよい。さて、スピン $1/2$ のパウリ演算子を $\{I, X, Y, Z\}$ と書くことにし、 n 量子ビットのパウリ演算子を $P_i \in \mathcal{P} = \{I, X, Y, Z\}^{\otimes n}$ を用いて、ハミルトニアン H は、

$$H = \sum_i h_i P_i, \quad (2)$$

の形で与えられているとする。ここで h_i は実数の係数である。VQE では、このハミルトニアン H のエネルギー期待値

$$\langle H(\theta) \rangle = \langle \psi(\theta) | H | \psi(\theta) \rangle, \quad (3)$$

を求めるために、それぞれの項 P_i のエネルギー期待値

$$\langle P_i(\theta) \rangle = \langle \psi(\theta) | P_i | \psi(\theta) \rangle, \quad (4)$$

を別個に測定する。例えば、 $P_i = X \otimes Z \otimes I \otimes \dots \otimes I$ であれば、1つ目と2つ目の量子ビットをそれぞれ X 基底と Z 基底で測定し、その測定結果 ± 1 の積の期待値を求めればよい。このように $\langle P_i \rangle$ の期待値は、個々の1量子ビットへの測定と測定結果の処理から取得できる。その後古典計算機上で

$$\langle H(\theta) \rangle = \sum_i h_i \langle P_i(\theta) \rangle, \quad (5)$$

を計算し、ある変分パラメータ θ におけるエネルギー期待値 $\langle H(\theta) \rangle$ を得るのである。測定したエネルギー期待値 $\langle H(\theta) \rangle$ はパラメータ θ の関数となっており、回路パラメータ θ を調節することを繰り返し、 $\langle H(\theta) \rangle$ を最小化することで、近似的な基底状態を得るのがVQEの基本的な考え方である。ここで、パウリ演算子の集合 \mathcal{P} は量子ビットの数に対し指数個の元を持つが、通常の物理系のハミルトニアンならば、そこには量子ビットの数に対し多項式個のパウリ演算子しか含まれないことに注意されたい。このためエネルギー期待値の計測にかかる時間は、量子ビットの数に対して指数的に増えることはない。

さて、このようなアルゴリズムにおいてその性能の鍵となるのは、回路パラメータ θ の最適化方法であろう。ある関数 (VQE の場合は $\langle H(\theta) \rangle$) を最小化するパラメータを求めよ、というパラメータの最適化問題が与えられたとき、その解法は、目的関数の勾配情報を使うものと使わないものと大きく二分できる。一般には、勾配情報を使う勾配降下法などのほうが勾配情報を使わない方法と比べて有利だと考えられる。そこで筆者らは、エネルギー期待値 $\langle H(\theta) \rangle$ の θ に関する勾配を求めるための手法を提案している。¹⁰⁾ 試行状態のパラメータは、1量子ビットの回転ゲート、例えば $e^{-i\theta X/2}$ のような形で与えることが多い。これは量子ビット同士の相互作用を必要とする2量子ビットゲートのほうが制御が難しいためである。このようなとき、

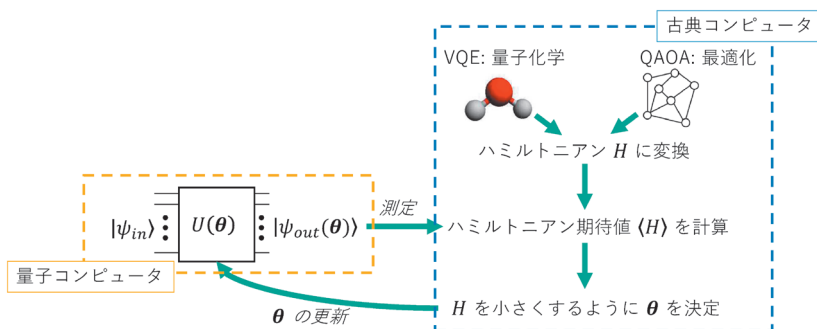


図1 変分量子アルゴリズムの代表例, Variational Quantum Eigensolver と Quantum Approximate Optimization Algorithm. 与えられたハミルトニアン H の基底状態を探索する。

ハミルトニアン H の期待値は

$$\langle H(\theta) \rangle = \langle 0 | e^{i\theta X} H e^{-i\theta X} | 0 \rangle \quad (6)$$

と書けるので、その微分は

$$\frac{d\langle H(\theta) \rangle}{d\theta} = \frac{i}{2} \left(\langle 0 | X e^{i\theta X/2} H e^{-i\theta X/2} | 0 \rangle - \langle 0 | e^{i\theta X/2} H e^{-i\theta X/2} X | 0 \rangle \right) \quad (7)$$

である。この物理量はそのままでは観測することが難しい。この微分の表式を直接評価する手法としては、補助ビットを用いて間接測定することも考案されている¹¹⁾が、NISQデバイスの限られたリソースの中で1量子ビットを消費するのは大きな痛手である。しかしながら実はこの量は

$$e^{-i\theta X/2} = \cos \frac{\theta}{2} I - i \sin \frac{\theta}{2} X \quad (8)$$

というパウリ演算子のよく知られた性質を用いると、

$$\frac{d\langle H(\theta) \rangle}{d\theta} = \frac{1}{2} \left(\left\langle H \left(\theta + \frac{\pi}{2} \right) \right\rangle - \left\langle H \left(\theta - \frac{\pi}{2} \right) \right\rangle \right) \quad (9)$$

のように、補助量子ビットを使わずとも評価できるのである。この手法は、 θ の値を微妙に変化させて勾配を評価する単純な差分法に似ているが、それよりもはるかに誤差に対して堅牢である。

VQEが古典コンピュータに対して優位性があるかどうかはまだ分かっていない。そもそも、与えられたハミルトニアンの基底状態とそのエネルギーを求める問題は、最悪の場合量子コンピュータがあっても効率よく解けないQMA (Quantum Merlin Arthur問題と呼ばれ、NP問題の量子版と思ってもらってよい)⁴⁾という非常に難しいクラスに属する。例えば、ランダムネスのある2次元ハバードモデルの基底エネルギーを求める問題は、どのQMA問題よりも難しい(QMA困難)問題であり、量子コンピュータを用いても効率よく解けないと考えられている。しかしながら、我々が興味をもっているような自然界で実現しているような多くの系であれば(自然界で実現しているのだから)、量子力学と互換性のある量子コンピュータでその状態を効率よく生成できると期待されている。また、ノイズがない場合には、より広いヒルベルト空間(よりエンタングルした状態)を探索できるので、古典コンピュータによる変分法に対して優位性があると期待される。ノイズがあり、量子回路のサイズが制限された場合に従来手法に対する優位性があるかどうかは現在研究が進められているところである。例えば、VQEの提案論文⁷⁾で用いられたUnitary Coupled Clusterと呼ばれる試行状態は、古典コンピュータでは効率的に生成することができない状態である。

VQEに関して最近の研究をいくつか紹介しよう。一つの方向性は、試行状態を生成するパラメータ付き量子回路 $U(\theta)$ の構成に関する研究である。ノイズの影響を抑えるために使われる量子ゲートを可能な限り少なくしたり¹²⁾、そもそもノイズの影響を受けにくい量子回路を構成したり¹³⁾といった研究がある。回路パラメータ θ の最適化手

法を改善することも一つの大きなテーマである。例えば、従来の変分的アプローチでも使われているような虚時間発展によるパラメータ最適化¹⁴⁾が提案されている。また、基底状態だけではなく、励起状態を得るためにも様々なアプローチがなされている。¹⁵⁻¹⁷⁾我々も一方式として、量子回路のユニタリ性を有効に利用して低エネルギーの部分空間を求める、部分空間探索VQE (subspace search VQE)を提案している。¹⁸⁾

VQEの枠組みは0,1変数からなる最適化問題(0,1計画問題)にも応用できる。最小化するハミルトニアンとして、いわゆるイジングハミルトニアン

$$H = \sum_{i=1}^n h_i Z_i + \sum_{i=1}^n \sum_{j=1}^i J_{ij} Z_i Z_j \quad (10)$$

を採用してVQEを適用すれば良い。イジングハミルトニアンのような最適化問題にVQEを適用したとき、特に、量子近似最適化法¹⁹⁾(Quantum Approximate Optimization Algorithm, QAOA)とも呼ばれており、量子アニーリングの拡張として位置付けられている。アメリカのベンチャーであるRigetti社は、超伝導量子ビットを使ったQAOAの実験を発表している。²⁰⁾

3. 機械学習への応用: 量子回路学習

試行関数のパラメータを変分的に最適化することは物理だけでなく、幅広い分野で用いられている。その最たる例が、ニューラルネットワークによる機械学習であろう。以降では、機械学習の基本的な考えから解説し、量子コンピュータを用いた機械学習、量子機械学習について解説する。特に、前節で紹介した量子古典ハイブリッド型の変分アルゴリズムであるVQEの考え方を機械学習へと拡張した量子回路学習と、その応用について解説する。

3.1 量子機械学習

機械学習とは、過去に得られている多数のデータ $\{x_i\}$ からモデル $f(x)$ を構築し、新たなデータ x に対する予測を行う枠組みである。一般に、そのようなモデル $f(x)$ を構築することを“学習”と呼ぶ。 $f(x)$ の形には様々なものが用いられる。例えば線形回帰では

$$f(x) = w \cdot x + b, \quad (11)$$

(w, b は実数のパラメータ)を仮定するし、ロジスティック回帰では

$$f(x) = e^{w \cdot x + b} / (1 + e^{w \cdot x + b}) \quad (12)$$

を、近年話題のディープラーニング(多層ニューラルネットワーク)では、

$$f(x) = g_n(W_n(\dots(g_2(W_2 g_1(W_1 x))\dots)) \quad (13)$$

(W_i はパラメータ行列、 g_n は活性化関数)のような階層的な形を用いる。

機械学習では、特徴量空間という概念も非常に重要であ

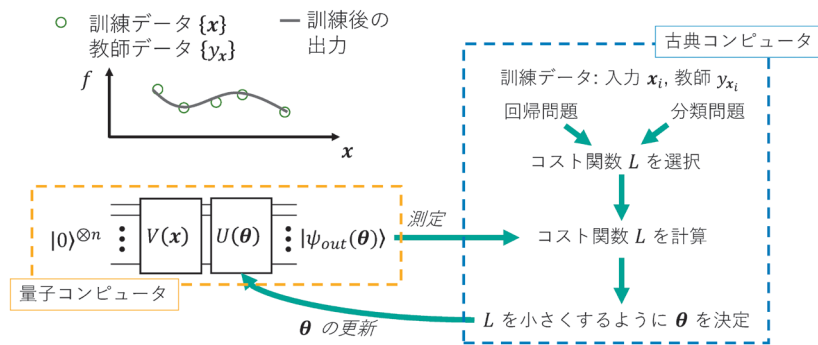


図2 量子回路学習 (QCL) の概念図.

る. 通常データ $\{x_i\}$ として与えられるのは, 画像のピクセル値を羅列したものや音声の波形データなど, そのままでは意味の分かりにくいものである. 例えば音声認識であれば, 与えられたデータに対してフーリエ変換を施し周波数成分を取り出したもののほうが, そのデータの特徴をつかみやすいだろう. このように, 与えられたデータ $\{x_i\}$ に対して, 特徴的な量を抽出する目的で何らかの処理を施した量 $\{\phi(x)\}$ を特徴量と呼び, その量が存在する空間を特徴量空間という. 特に分類問題では, 元データ $\{x_i\}$ ではうまく分類することができないとき, データを元の次元よりも高い次元を持つ空間へ持っていくといったことがよく行われている.

機械学習は大きく教師あり学習と教師なし学習, それと強化学習に分けられる. ここでは教師あり学習を例として, 学習がどのように行われるのかを説明しよう. 教師あり学習では, データ $\{x_i\}$ と対応する教師データ $\{y_i\}$ が与えられる. 例えば, $\{x_i\}$ は 0~9 の手書き数字の画像であり, 教師データ $\{y_i\}$ はそれぞれのラベル (0~9) である. このような教師あり学習では, 教師データ $\{y_i\}$ とモデル $\{f(x_i)\}$ の間の差を, モデル内のパラメータ (上の例では w, b や W_i) を調節することにより, できるだけ小さくすることで学習が行われる. $\{y_i\}$ と $\{f(x_i)\}$ の差を測る関数

$$\mathcal{L}(\{y_i\}, \{f(x_i)\}) \quad (14)$$

はコスト関数と呼ばれる. コスト関数の例としては, 回帰問題に使われる二乗誤差や, 分類問題に使われるクロスエントロピーなどがあげられるだろう. 教師あり/教師なし学習・強化学習のどれであっても, 行いたいタスクに適したコスト関数, もしくはそれに準ずるものを考える. コスト関数を最小化するようなモデル $f(x)$ をパラメータの最適化によって構築するのが機械学習である. さらなる詳細は文献 21 をあたられたい.

量子コンピュータを用いた機械学習は, 2009 年の Harrow-Hassidim-Lloyd による行列計算アルゴリズム²²⁾ (HHL アルゴリズム) に端を発し, 様々なアルゴリズムが提案されている. HHL アルゴリズムは適当な条件下で, $N \times N$ の行列 A に対する線型方程式 $Av = b$ を, $O(\log N)$ の時間で解くアルゴリズムである. これを応用し, 例えば量子サポー

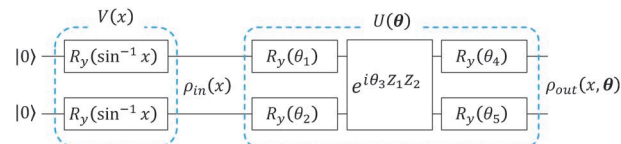


図3 2量子ビットを使ったQCLの例.

トベクターマシン²³⁾ や量子線形回帰^{24, 25)} などが作られている. これらのアルゴリズムは量子加速が示されているものの, 非常に多くの量子ゲートが必要であり, 本稿の主題であるNISQデバイスにとってはほとんど不可能である.

3.2 NISQデバイスによる量子機械学習

我々は, 変分的なアプローチによるNISQデバイス向けの量子機械学習アルゴリズムである, 量子回路学習¹⁰⁾ (Quantum Circuit Learning, QCL) を提案した. これは, 状態の最適化を目的とするVQEを拡張し, 量子回路からの出力と教師データとの誤差を最小化を目指して回路パラメータを調整するという提案である. 著者の一人は, この提案に先立って, 複雑な量子多体系の実時間ダイナミクスと線形回帰によって学習を行う, 量子レザバー計算を提案しており²⁶⁾ NMR固体量子スピンアンサンブル系を用いた実証実験²⁷⁾ も行っている. 量子レザバー計算は, 量子多体系の自然なダイナミクスをそのまま利用し, 線形回帰によってモデルを構築する方法であるのに対して, QCLは通常のニューラルネットワークのようにパラメータを調整することによってモデルを構築する. VQEにおいては古典コンピュータでは作り出すことのできない量子状態を試行関数とするところに量子コンピュータを利用する強みがあった. QCLは, テンソル積によって指数的に次元が大きくなるヒルベルト空間を特徴量空間として利用し, そこに作用するユニタリ変換のパラメータを調整することによってモデルを学習しようという提案である. つまり, ニューラルネットワークが線形変換のパラメータと与えられた非線形関数 (活性化関数) からモデルを構築するのに対して, 量子回路学習は, テンソル積構造に起因する非線形性とパラメータ付きユニタリ変換の調整によってモデルを学習することになる.

もう少し詳しくQCLについて解説していこう. 概念図を図2に, 2量子ビットを使った具体例を図3に示す.

QCLではまず、入力量子回路 $V(x)$ によって量子状態にデータ x をエンコードする。図3では一次元のデータ x によって決まる y 軸回転ゲート

$$R_y(\sin^{-1}x) = \sqrt{1-x^2}I - ixY \quad (15)$$

をそれぞれの量子ビットにかけている部分が入力量子回路に相当する。このゲートの直後の状態を表す密度演算子は

$$\rho_{\text{in}}(x) = \left(\frac{I + xX + \sqrt{1-x^2}Z}{2} \right)^{\otimes 2} \quad (16)$$

となっている。回転ゲートの引数を $\sin^{-1}x$ としているのは、一つ一つの量子ビットを測定したときに x に関して線形な出力を得るためである。この入力状態について、 $X \otimes X$ というオブザーバブルの期待値を測定すれば

$$\begin{aligned} \langle X \otimes X \rangle &= \text{Tr}(\rho_{\text{in}} X \otimes X) \\ &= x^2 \end{aligned} \quad (17)$$

となり、一つの量子ビットを見ていた時には表れない x^2 という関数を引き出すことができる。このようにたとえ $V(x)$ が非常にシンプルなゲート、例えば単一量子ビット回転を組み合わせたものであっても、量子系の直積構造から、量子状態

$$V(x)|0\rangle^{\otimes n} = \sum_k a_k(x)|k\rangle \quad (18)$$

の振幅 $a_k(x)$ はデータ x について高次の非線形性を持つ。このようにして生まれる指数的な次元の特徴量空間が、量子コンピュータを使う強みである。もちろん単純化のために x をそのまま入力したが、何らかの非線形関数 $\phi(x)$ として入力してもよい。

入力 $V(x)$ の後、適当な量子回路 $U(\theta)$ によって量子状態を変換し、出力状態 $\rho_{\text{out}}(x, \theta)$ を得る。QCLではこの出力状態の適当なオブザーバブル O の期待値 $\langle O(\theta, x) \rangle$ を測り、これを量子回路が構築するモデル $f(x|\theta)$ と考える。図3の例で、例えばオブザーバブルとして第一量子ビットの Z 演算子をとる、

$$f(x|\theta) = \text{Tr}(\rho_{\text{out}}(x, \theta) Z \otimes I) \quad (19)$$

としてみよう。 $U(\theta)$ のパラメータを適切に調整すれば、このモデル $f(x|\theta)$ は

$$f(x|\theta) = x^2 \quad (20)$$

とも、

$$f(x|\theta) = x \quad (21)$$

とも、またこれらの関数の線型結合

$$f(x|\theta) = c_1x + c_2x^2 \quad (22)$$

ともなりうるだろう。さらに多くの量子ビットを使えば、これらの関数の数は指数的に増えていく。QCLは、このように様々な関数を表現できるモデル $f(x|\theta)$ を用いて、教師データ y_x を可能な限り再現するモデルが得られるよ

うに、パラメータ θ の最適化を行う枠組みである。

実はQCLでは、量子回路がユニタリ変換であることにより、過学習をある程度抑えられるという効果がある。一般に過学習とは、モデル f にデータ点数と同程度の数の学習パラメータを設定してしまった時に発生する現象である。モデルが過学習を起こしてしまうと、訓練データ点“だけ”を完全に記憶してしまい、未知のデータが与えられたときに全く正しい予想ができなくなる。例えば線形回帰(式(11))では、過学習を起こすと線形結合の重み w のノルムが大きくなることが多い。そこでこれを回避するために、コスト関数に $\|w\|$ を足すことによってパラメータのノルムが大きくなりすぎることを抑制する、正則化と呼ばれる手法がある。翻ってQCLについて考えてみると、例えば式(22)における線型結合の重み c_1, c_2 は量子回路のユニタリ性によって $|c_1|^2 + |c_2|^2 = 1$ である。常にパラメータのノルムが1に制限されているため、正則化が自動的に行われているような形となっている。このことから、通常と同じ特徴量空間を使った線形回帰に比べれば、QCLは過学習を起こしにくいといえるだろう。

上記の関数の汎化だけではなく、分類問題もQCLで実行することができ、簡単な2値分類問題について論文で報告している。そして、このQCLによる分類を提案した直後、IBMのグループが超伝導量子ビットを用いて実験的にこれを実現している。²⁸⁾ それ以降も、NISQデバイスを持ちいた変分アプローチによる量子機械学習については、様々な提案が次々に発表されている。例えば、機械学習における生成モデルをNISQデバイスへと拡張し、これを量子コンピュータの性能評価へと応用する試みがある。²⁹⁾ また、カナダのベンチャー企業であるXanaduのグループが、NISQデバイス上での敵対的生成ネットワークの構築を提案している。³⁰⁾ これらの試みが、実際のNISQデバイス上でどれほどの性能を出すことができるかはいまだ未知数であるが、今後の発展に期待したい分野である。

3.3 量子データのための量子機械学習

以上では古典的なデータを量子コンピュータでどのように学習するか、という枠組みで議論を行ってきた。しかし、量子機械学習がその恩恵を一番受けることができるのは、おそらく量子的なデータやモデルの背景に量子系があるような場面の学習問題に適用したときであると考えられる。近年の研究で古典コンピュータによる量子系の学習が成果をあげている³¹⁾ 以上、量子コンピュータを量子系の学習に用いれば、さらなる成果が期待できるだろう。この文脈における研究としては、量子自己符号化器³²⁾ (quantum autoencoder) の提案は、量子系を量子コンピュータで学習しようという新しい取り組みである。量子自己符号化器の後続研究として、変分量子対角化³³⁾ (variational quantum diagonalization) がごく最近提案されている。我々も、量子回路学習を進展させ、量子コンピュータを用いて量子系を学習しようという方向の研究を進めている。

ここでは、VQEによって得られた基底状態を機械学習と同じように「汎化」させるという、最近の我々の研究について少し解説する。³⁴⁾ VQEは、一つのハミルトニアン H の基底状態を求めるアルゴリズムであったが、量子化学においては、例えば結合長によるエネルギーの変化を求めたいという要求がある。この目的を従来のVQEによって達成するには、それぞれの結合長 r におけるハミルトニアン $H(r)$ を与え、一つ一つの基底状態・基底エネルギーを求めてゆくほかなかった。一方で機械学習によってこのような問題を解く場合、ある少数の $\{r_i\}$ について学習を行い、未知の r が与えられた時も正しいエネルギーを返すようなモデルを構築するのが常道である。このように未知データに対しても正しい答えを返すことをモデルの汎化と呼ぶ。もしVQEの出力である基底状態をこの意味で汎化させられれば、調べたい系のパラメータについてその系がどのように変化するか、比較的少数の点において量子回路を訓練することで調べられるだろう。そこで我々は、回路深さを無限大とした極限で断熱時間発展を必ず実現できるように、パラメータ数の比較的少ない量子回路 $U(\theta)$ を提案し、VQEの出力の汎化を試みた。そのような回路を選んだのは、十分長い時間をかけた断熱時間発展を行えば、目的のハミルトニアンの形によらずに基底状態が得られるという事実に立脚したいと考えたためだ。もちろん、回路深さが有限であるときには、断熱時間発展を厳密に行うことはできないため、各結合長 r において $U(\theta)|0\rangle$ をVQEによって最適化したとき、得られる最適な回路パラメータ θ^* は r に依存した形 $\theta^*(r)$ として得られる。我々は、少数の点 $\{r_i\}$ (訓練点) においてVQEを走らせて $\{\theta^*(r_i)\}$ を求め、訓練点の間における $\theta^*(r)$ を、訓練点における $\{\theta^*(r_i)\}$ の補間によって求めるという単純なアプローチによって、最適なパラメータ $\theta^*(r)$ を学習することを提案した。このアプローチにより、物理系のパラメータについてその性質がどのように変化するか、従来の単純なVQEよりも高速に予測することができると考えている。また、磁場の強度などのパラメータについて、量子相転移現象などを量子回路で学習し基底状態を汎化することができるか、といった興味深い問題も残されている。

3.4 量子最適制御のための量子機械学習

量子データの機械学習のみならず、近年では量子系の精密な制御にも機械学習的な手法が使われ始めている。量子系の制御は核磁気共鳴の分野で古くから研究がなされているが、最近になって、特に量子コンピュータの実現に向けてさらに高度な手法が提案されている。その皮切りとなったのは、最急上昇法によるパルスエンジニアリング³⁵⁾であり、今では制御したい量子系をそのまま用いて量子制御を精緻化する手法が現れている。^{36,37)} 筆者の一人は、超伝導量子ビットの実験グループと共同で、超伝導量子ビットに対する2量子ビット演算の精度の向上のために、量子回路による変分アプローチを適用する方法 (VQGO: Varia-

tional Quantum Gate Optimization) を提案している。³⁸⁾ これは、1量子ビット演算の精度が99.9%以上という高精度であることに比べ、2量子ビット演算は、高次の摂動項やパルスの混信などの寄与によって、精度が低くなるということに着目している。これまで、これらの効果を打ち消すパルスが考案されてきたが、我々の変分アプローチでは、そのような量子操作を最適化によって見つけようというものである。また、VQEが状態の最適化であるのに対し、VQGOは量子操作の最適化であり、より一般的な変分アルゴリズムへの拡張と捉えることもできる。

4. ノイズ低減のための取り組み

ここまではNISQデバイスをどのように応用していくか、そのアプリケーションを紹介してきた。だが、エラーが無視できない実際のNISQデバイスを使用するには、ノイズの影響をどれだけ減らせるかが鍵となる。もちろんノイズそれ自体を減らすには、ハードウェア側の性能向上を待つしかない。それではソフトウェア側でやれることはないのだろうか？ 決してそのようなことはなく、測定結果の統計処理をうまく行ったり、量子回路の構成を工夫したりすることで、測定される観測量に対するノイズの影響を少なくしようという研究がなされている。これらの研究で提案されている手法は、一般にノイズ補償 (error mitigation) 手法と呼ばれる。

一つの手法は、量子ビットにかかるノイズの大きさを人為的に変化させ、それに伴う観測量の変化から、理想的な観測量を復元するというものである。³⁹⁻⁴¹⁾ この手法は外挿法によるノイズ補償 (error mitigation with extrapolation) と呼ばれている。観測する量を O 、ノイズの大きさが ϵ であるときの O の期待値を $\langle O(\epsilon) \rangle$ としよう。外挿法では、人為的にノイズを変化させた系における期待値 $\{\langle O(\epsilon_i) \rangle\}$ を数点測定する。その後、 ϵ について線型関数 $a\epsilon + b$ もしくは指数関数 $Ae^{-a\epsilon}$ などによってフィッティングを行い、その外挿によってノイズ無しの系における理想的な期待値 $\langle O \rangle_{\text{ideal}} = \langle O(0) \rangle$ を得るのである。この手法は、実験的にもその有効性が検証されている。⁴²⁾ ここでは詳細は述べないが、ノイズモデルを知った上でその補償を行う、擬確率分布を用いた手法も提案されている。^{39,41)}

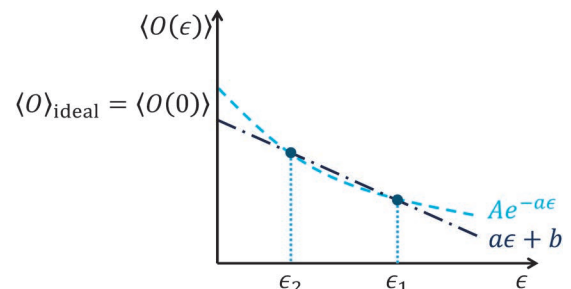


図4 外挿法による誤差緩和。 ϵ はノイズの大きさ、 $\langle O(\epsilon) \rangle$ はあるノイズの大きさにおける観測量 O の期待値。

5. おわりに

本稿では、すでに実現している、もしくは近未来的に実現する規模の量子コンピュータを用いた量子古典ハイブリッド変分アルゴリズムについて解説した。量子多体系の基底状態を探索するVQE(変分量子固有値法)は、パラメータ付き量子回路によって生成される量子状態を試行関数として用いる変分法であった。これは、従来コンピュータ上で実行することが前提となっている、密度行列繰り込み群やテンソルネットワークなどの変分アプローチの量子コンピュータへの拡張と捉えることができる。基底状態だけでなく、励起状態も変分アルゴリズムとして求めることができることも述べた。Feynmanが指摘したように、量子コンピュータ上では時間発展をそのままシミュレーションすることができるので、基底状態と励起状態の間の複雑な遷移ダイナミクスも取り込むことができるだろう。また、エネルギー固有状態などの状態の最適化だけではなく、入力データと教師データから量子回路をモデルとして学習、量子回路学習も紹介した。

これら、NISQデバイスをターゲットにした量子アルゴリズムでは、できるだけ小規模の量子コンピュータで多少ノイズがあっても動作するように構成されている。このような制約のもと機能を低下させないために、NISQデバイス用の量子アルゴリズムでは、様々なアイデアが試されている。このように、NISQデバイスの登場によって、大規模な量子コンピュータだけを主眼としていた頃の量子アルゴリズム設計とは異なった方向性で新しいアイデアが提案されている。まさに、クロック数もメモリ容量も小さかったファミリーコンピュータ上でのゲーム開発(スーパーマリオブラザーズやドラクエなど)に似た面白さがある。そして、ここで生まれた創意工夫は、NISQ時代だけではなく、大規模な量子コンピュータが登場した後も役立つことだろう。

最近、ニューラルネットワークを用いた量子状態の表現や物質相の分類、そして深層学習と物理の接点など、機械学習分野のアイデアがますます物理に浸透してきている。多体量子系を学習するには、量子コンピュータ上でモデルを構築し学習する量子機械学習が有利であると期待される。VQEは、量子コンピュータ上でモデル(試行状態)を構築し変分的に最適化する方法であった。また、量子回路学習は、ニューラルネットワークの代わりに量子系のテンソル積構造と巨大な次元のユニタリ変換を利用する、という試みである。この、多体量子系をモデル化する試行関数、機械学習、量子計算の距離が急速に縮まっているように感じる。今後の展開に目が離せない。

参考文献

- 1) 阿部英介, 伊藤公平, 応用物理 **86**, 453 (2017).
- 2) J. Preskill, Quantum **2**, 79 (2018).
- 3) S. Boixo et al., Nat. Phys. **14**, 595 (2018).
- 4) 森前智行, 日本物理学会誌 **74**, 98 (2019).
- 5) 西野友年, 奥西巧一, 引原俊哉, 物性研究 **68**, 133 (1997).

- 6) 原田健自, 数理解析研究所講究録 **1848**, 83 (2013).
- 7) A. Peruzzo et al., Nat. Commun. **5**, 4213 (2014).
- 8) B. Bauer et al., Phys. Rev. X **6**, 031045 (2016).
- 9) A. Kandala et al., Nature **549**, 242 (2017).
- 10) K. Mitarai et al., Phys. Rev. A **98**, 032309 (2018).
- 11) G. G. Guerreschi and M. Smelyanskiy, arXiv:1701.01450 (2017).
- 12) P.-L. Dallaire-Demers et al., arXiv:1801.01053 (2018).
- 13) I. H. Kim and B. Swingle, arXiv:1711.07500 (2017).
- 14) S. McArdle et al., arXiv:1804.03023 (2018).
- 15) J. R. McClean et al., Phys. Rev. A **95**, 042308 (2017).
- 16) O. Higgott, D. Wang, and S. Brierley, Quantum, **3**, 156 (2019).
- 17) S. Endo et al., Phys. Rev. A **99**, 062304 (2019).
- 18) K. M. Nakanishi, K. Mitarai, and K. Fujii, arXiv:1810.09434 (2018).
- 19) E. Farhi, J. Goldstone, and S. Gutmann, arXiv:1411.4028 (2014).
- 20) J. S. Otterbach et al., arXiv:1712.05771 (2017).
- 21) C. M. Bishop, *Pattern Recognition and Machine Learning* (Springer, 2006).
- 22) A. W. Harrow, A. Hassidim, and S. Lloyd, Phys. Rev. Lett. **103**, 150502 (2009).
- 23) P. Rebentrost, M. Mohseni, and S. Lloyd, Phys. Rev. Lett. **113**, 130503 (2014).
- 24) N. Wiebe, D. Braun, and S. Lloyd, Phys. Rev. Lett. **109**, 050505 (2012).
- 25) M. Schuld, I. Sinayskiy, and F. Petruccione, Phys. Rev. A **94**, 022342 (2016).
- 26) K. Fujii and K. Nakajima, Phys. Rev. Appl. **8**, 024030 (2017).
- 27) M. Negoro et al., arXiv:1806.10910 (2018).
- 28) V. Havlicek et al., Nature **567**, 209 (2019).
- 29) M. Benedetti et al., npj Quant. Inf. **5**, 45 (2019).
- 30) P.-L. Dallaire-Demers and N. Killoran, Phys. Rev. A **98**, 012324 (2018).
- 31) G. Carleo and M. Troyer, Science **355**, 602 (2017).
- 32) J. Romero, J. P. Olson, and A. Aspuru-Guzik, Quantum Sci. Technol. **2**, 045001 (2017).
- 33) R. LaRose et al., npj Quant. Inf. **5**, 8 (2019).
- 34) K. Mitarai, T. Yan, and K. Fujii, Phys. Rev. Appl. **11**, 044087 (2019).
- 35) N. Khaneja et al., J. Magn. Reson. **172**, 296 (2005).
- 36) J. Li, X. Yang, X. Peng, and C.-P. Sun, Phys. Rev. Lett. **118**, 150503 (2017).
- 37) D. Lu et al., npj Quant. Inf. **3**, 45 (2017).
- 38) K. Heya et al., arXiv:1810.12745 (2018).
- 39) K. Temme, S. Bravyi, and J. M. Gambetta, Phys. Rev. Lett. **119**, 180509 (2017).
- 40) Y. Li and S. C. Benjamin, Phys. Rev. X **7**, 021050 (2017).
- 41) S. Endo, S. C. Benjamin, and Y. Li, Phys. Rev. X **8**, 31027 (2018).
- 42) A. Kandala et al., Nature **567**, 491 (2018).

著者紹介

御手洗光祐氏: 量子計算, 特に量子コンピュータをどのように実用していくかに興味がある。

藤井啓祐氏: 専門は量子情報, 特に量子計算, 量子誤り訂正, そして量子計算と物理との接点に興味がある。

(2018年12月15日原稿受付)

Variational Algorithms on a Quantum Computer and Their Applications for Machine Learning

Kosuke Mitarai and Keisuke Fujii

abstract: The recent progress of the hardware of the quantum computer are dramatic, and quantum computers with several hundreds of qubits might be within reach in a few years. This progress is motivating the world to search for applications of such devices. Here we give the overview of the recent development of algorithms for near-term, noisy quantum computers. This article is focused especially on so-called variational quantum algorithms where classical and quantum computers works together to solve a given task. The near-term quantum computers are expected to be applied in the area ranging from machine learning to quantum chemistry calculation. We also review some of the approaches to mitigate the noise effect.