連載

5分で分かる!? 有名論文ナナメ読み

曽我部東馬（電気通信大学）

C-C. Chen, M. Watabe, K. Shiba, M. Sogabe, K. Sakamoto and T. Sogabe :
On The Expressibility and Overfitting of Quantum Circuit Learning

ACM Transactions on Quantum Computing, Vol.2, No.8, pp.1-24 (2021)

過学習，PAC 学習理理論

本稿の紹介を始める前に，背景知識となる学習理論，過学習そして VC 次元について簡単にふれたい．機械学習は常に，過学習と呼ばれる潜在的な危険性がある．この場合，学習モデルはデータのノイズにさえ適合してしまい，未知のデータに対する予測精度が著しく低下する現象である．それは知識を本当に理解せずに「丸暗記」して学習する学生のようなもので，未知のデータに対する対応能力が喪失する．未知の問題に対する汎化機能の獲得は機械学習分野の中心タスクである．

過学習を扱う研究手法や理論は多数開発されている．その中で代表的な理論は 1970 年代に Vapnik and Chervonenkis によって提案された VC 次元理論である[1]．VC 次元より以前に，ヘフディングの不等式を拡張した Union Bound という汎化誤差を評価する基準が提案されているが，学習モデルの集合サイズ $|\mathcal{H}|$ が有限の場合しか適用できないという制限があった．それに対して，VC 次元理論は無限の集合サイズ $|\mathcal{H}|$ でも学習モデルの汎化誤差の上界を定量的に解析でき，機械学習理論の発展に大きく貢献した．また，Union Bound 学習理論も VC 次元理論も基本的には，汎化誤差を訓練誤差を通して確率的に評価する「近似的に正しい学習：PAC（Probably Approximately Correct）学習」の仕組みとなっている[2]．

VC 次元理論の概略

VC 次元の理論背景と詳しい導出は紙面の都合上，省略するが，図 -1 の例を通して要点だけ説明する．3 つの 2 次元データ赤と青の 2 値分類に対して，線形分離器 $h(x)=sign(\boldsymbol{a}^T\boldsymbol{x}+b),\boldsymbol{a} \in R^2;\ b \in R$ を適用すると仮定する．a と b の値は連続的に変化することが可能なので，図 -1（a）のように，分類できる学習モデルの集合サイズは $|\mathcal{H}| \rightarrow \infty$ になってしまい，従来の Union Bound 学習理論が適用できなくなる．そこで，VC 次元理論は，単純に学習モデルの数を数えるのではなく，学習器が分類できるデータの数を元に集合サイズ $|\mathcal{H}|$ を定義するという斬新な理論体系となっている．VC 次元理論の VC 次元 d_{VC} は堅く定義すると次のようになる：「仮説集合 $\mathcal{H}$ の仮説 h による $(x_1,\cdots,x_n)$ のラベルのセット $(h(x_1),\cdots h(x_n))$ の成長関数 Growth (H, n) を 2^n としたときの最大のサンプル数 n を VC 次元 $d_{VC}=n$ とする」．ちょっと分かりにくい概念なので，例を通して簡単に説明する．図 -1（b）に示すように，3 つのデータの赤と青の組合せパターンは 8 となる．図 -1（b）の左側に示している 3 つのデータの位置関係であれば，図 -1（a）に示した線形分離学習

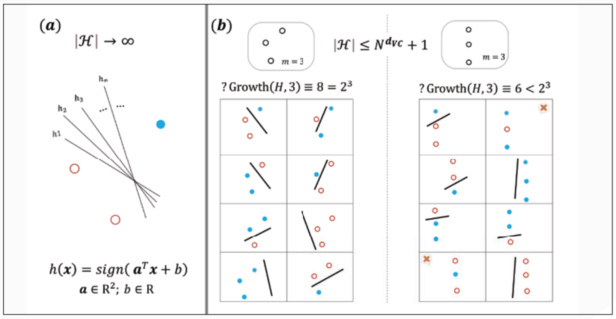

図 -1　VC 次元の模式図　(a) 従来の Union Bound 理論における線形分離学習モデルの集合サイズ $|\mathcal{H}|$ が発散，(b) VC 次元による線形分離学習モデルの集合サイズ $|\mathcal{H}|$ が上界（N^{dvc}+1）を持つ

器を使用した場合，8パターンに分離できることは図からすぐ検証できる．VC次元の定義から2変数の線形分離学習器の成長関数 Growth(H, 3) は最低でも $2^3=8$ となる可能性が示される．また，定義の中に最大のサンプル数 n という制約条件があるので，図-1（b）の右側の6パターンにしか分離できない3つのデータの位置関係の場合は，成長関数として見なさない．しかし，分離できる最大サンプル数 n の上限が果たして $n=3$ 以上あるのかどうかは不明である．実はこれがVC次元の理論を活用する上での最大の難所といえる．VC次元 d_{VC} の正確な値は数理的にしか証明できないので，複雑な学習モデルを用いた場合，d_{VC} を決定するのは非常に困難である．そういう難点があるものの，VC次元に基づいた学習理論は，ほかの学習理論にないいくつかの利点がある．(1) 分離できる最大サンプルを元にした理論であるため，入力データの分布に依存しない；(2) 漸近的ではなく小さなサンプルサイズにも適用できる；(3) 学習モデル集合の特性を表しているので，個々のモデルから独立している；(4) ノイズの多いデータに対しても機能する．これらの利点は，最新の機械学習にとって重要で，VC理論は，後に人工ニューラルネットワークに適用され，さらに一世を風靡したサポートベクタマシン（SVM）のアルゴリズムの中核にもなったのである．

以上，本論文の内容を理解するための過学習や汎化誤差の上界，VC次元理論，そしてVC次元の計算の困難さとその応用上の利点に関する背景知識を紹介した．今回紹介する論文の中心テーマである「量子コンピュータを用いた量子機械学習」のようなまだ黎明期の最先端研究分野においては，VC次元の計算が容易ではないので，VC次元理論を適用した量子機械学習モデルの汎化誤差評価に関する研究はほぼなかった．

量子VC次元の確立とテンソルネットワーク理論

この論文で著者らはVC理論と量子回路学習の関係の確立を試みた．Control Z-gate Hardware Efficient Ansatz（CZ-HEA）という特定の量子回路学習モデルのVC次元は，有限の上界と下界を持つことを理論的に証明した．それによって，当量子回路学習モデルはPAC学習可能であることを世界ではじめて示した．また，VC次元に基づく量子回路学習モデルの過学習評価の正当性を考察するために，量子回路の説明能力を評価する情報理論基準KL-divergenceを用いて比較実験も行い，両者の一致性からVC次元の信頼性を左証した．ここで，本論文の研究成果に最も寄与したテンソルネットワーク理論に関連する内容を簡単に要約する．VC次元のより厳密な上界を導出するために，テンソルネットワーク理論のライトコーン制限を導入した．図-2は

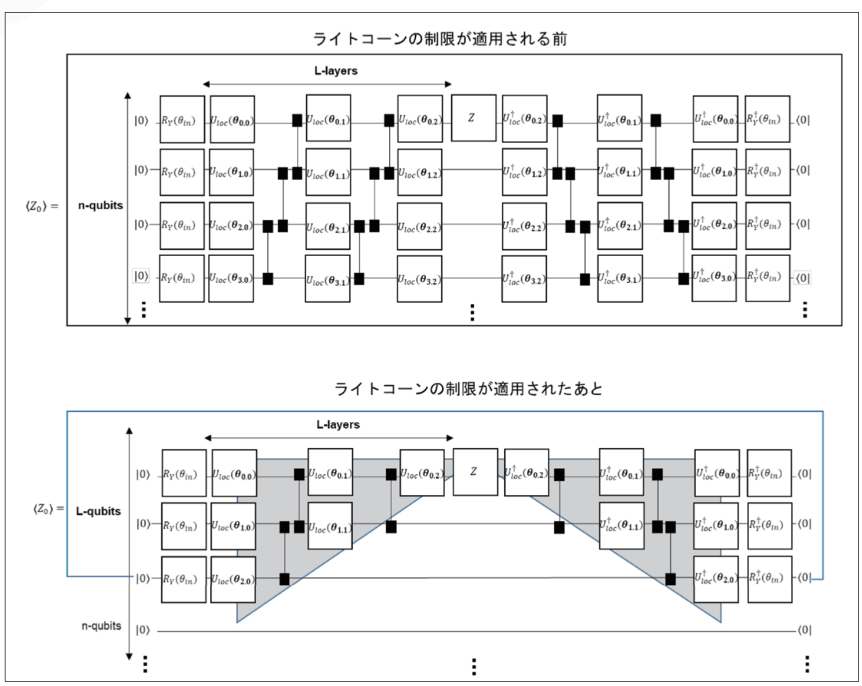

図-2　CZ-HEA量子回路におけるライトコーン制限の模式図

CZ-HEA という特定の量子回路を示している．ライトコーン制限を考慮しなければ，図 -2（a）に示しているように，d_{VC} は一般的に知られている上界

$$d_{VC} \leq (2\frac{n}{d} + 1)^{2d}$$

として与えることができる．ここでの d は入力データの次元あるいは（特徴の数）n は量子ビット数とそれぞれ対応している．しかし，この一般化した d_{VC} は量子ビット数が無限に増大するにつれ，汎化誤差が上界を持たなくなり，学習モデルの PAC 学習能力の保証を失ってしまう．それに対して，図 -2（b）に示しているようにライトコーン制限を考慮した場合は，量子ビット数 n は無限に増大しても，学習に寄与する有効な量子ビット数はライトコーン制限の元で，量子回路の深さ L に制限される．それ故，VC 次元は下記のような有限な上界を持つことが証明された．

$$d_{VC} \leq (2\min(\frac{n}{d}, \left\lfloor\frac{2L+1}{d}\right\rfloor + 1) + 1)^{2d}$$

量子機械学習は近年，量子コンピューティング分野においてホットトピックの 1 つとして，量子技術と最先端の AI 技術の融合による相乗効果が期待されている．量子回路学習モデルは，御手洗らによって提案され [3]，2018 年 IBM によって実験的に実証された [4]．本論文は学習理論から量子回路学習モデルの汎化誤差を世界で初めて厳密に評価できたので，量子機械学習の分野においては，まさに歴史に残る論文の 1 つとして位置づけられるかと思う．読者にもぜひご一読をお勧めしたい．

参考文献

1) Vapnik, V. N. and Chervonenkis, A. Ya. : On the Uniform Convergence of Relative Frequencies of Events to Their Probabilities. Theory of Probability & Its Applications 16, 2 (1971), 264-280, https://doi.org/10.1137/1116025
2) Valiant, L. G. : A theory of The Learnable. Commun ACM 27, 11 (1984), 1134-1142, https://doi.org/10.1145/1968.1972
3) Mitarai, K., Negoro, M., Kitagawa, M. and Fujii, K. : Quantum Circuit Learning, Phys Rev A 98 (2018), 032309, http://dx.doi.org/10.1103/PhysRevA.98.032309
4) Havlicek, V., Corcoles, A. D., Temme, K., A. Harrow, A. W., Kandala, A., Chow, J. M. and Gambetta, J. M. : Supervised Learning with Quantum-Enhanced Feature Spaces, Nature 567 (2019), 209-212, https://doi.org/10.1038/s41586-019-0980-2

（2021 年 11 月 25 日受付）

曽我部東馬（正会員）

sogabe@uec.ac.jp

電気通信大学 i- パワードエネルギーシステム研究センター准教授．先進エネルギーデバイスと AI 融合技術研究ステーション長．エネルギー創生・変換，量子アルゴリズムの開発と AI を用いた量子制御に関する研究に従事．