

# 機械学習，ニューラルネットワークの基本知識

岡野原 大輔

## 本連載のはじめに

化学の分野において，AIはさまざまなかたちで活用されるようになってきている。その代表例が2024年にノーベル化学賞を受賞したAlphaFoldである。これはAIを使うことでタンパク質の構造予測をはじめてこれまでにない精度で実現した。

本連載では，このような計算化学やシミュレーション領域と，ニューラルネットワークを中心としたAIの融合に焦点を当て，その最新事例をAI側からの立場で紹介していく。計算化学およびシミュレーションは，物理や化学の多様な現象を理論的に解き，分子や材料の性質を予測する分野である。この分野では，分子構造の最適化，反応経路の解析，材料設計など，多岐にわたる応用が行われている。一方で，AIは学習用データからその背後にある規則性や特徴を抽出し，それをを用いて未知のデータについても予測できるように進化してきた。たとえば，ある特徴をもつ分子が特定の物性を示しやすいといったような予測が可能である。

しかし，計算化学で使われるAIの事例では，必ずしも問題を支配するような方程式や法則を見つけることが究極の目標ではない。むしろ，汎化性を備えたモデルが問題の解を高精度かつ高速に予測できることが重要である（もちろん，人間にも解釈可能な単純で重要なルールを見つけることができる場合もある）。言い換えれば，現在のAI技術は，複雑な現象をそのままのかたちでその特性や現象を予測できる能力を提供しているといえる。

ディラックは，「化学のすべてと物理の大部分の問題に対して解くべき方程式はわかったが，問題はその方程式が複雑すぎて解けないことだ」という有名な言葉を残している。量子化学計算はこれを解くために発展し，長年にわたるさまざまな工夫やアイデアにより，多くの問題が解けるようになってきた。そこにAIという新しいツールが加

わり，さらに多くの問題を効率よく高精度に解くことが可能になりつつある。これにより，地球上に存在しないタンパク質や化合物についても，その構造や物性を実験を行わずに予測することが徐々にではあるが可能になってきた。今後予測技術が発展し，候補化合物がどのような反応を起こすのか，さらにはその制御方法についても予測できるようになるだろう。

一方で，「ブラックボックスで予測できたとしても科学になっていないのではないか」という批判もある。本連載の最後では，AI自体が物理法則や対称性などを見つけていく手法についても紹介する。

## 本連載の構成

本連載は，専門知識がない人でも読みやすい内容を目指している。興味をもたれた方は，本文中で紹介する出典をたどり，さらに深く学んでいただければ幸いである。

全12回を予定しており，以下のテーマについて順次紹介していく予定である。これらのトピックを通じて，トレンドを知ってもらうとともに，現時点での限界や今後の研究の方向性についても掘り下げていく。

なお，この第1回目は

第1回 | 連載内容の概要，機械学習やニューラルネットワークの基本知識

を紹介する。

## I. 高次元関数のモデル化

連載の前半では，量子化学計算にニューラルネットワークがどのように活用されているのかを紹介する。計算化学

やシミュレーションの分野では、時空間やそれに関連する関数をいかに表現するかが重要である。既知の対称性や特性を考慮しつつ、どのようにデータから学習するかについて解説する。

具体的には以下のテーマを取上げる予定である。

- 第2回 | ニューラルネットワークポテンシャル
- 第3回 | DFT（密度汎関数理論）における交換相関汎関数のニューラルモデル化
- 第4回 | 波動関数のニューラルモデル化

これらの手法はいずれも、ニューラルネットワークが高い表現力や情報圧縮能力を活用し、従来の手法では実現が難しかった効率化や高精度化を可能にしている。

## II. AIを使った予測・生成

次に、AIを活用した候補物質の生成、反応の予測、高速化について紹介する。この分野は、現在注目されている「生成AI」の技術と密接に関連する領域である。

具体的には以下のテーマを取上げる。

- 第5回 | 反応候補の予測・列挙
- 第6回 | 結晶構造予測および逆生成

また、タンパク質の構造予測や、平衡分布からのサンプリングについても取上げる。これらは生成AIの進展を象徴する事例である。

- 第7回 | タンパク質の構造予測（AlphaFold）
- 第8回 | 生成モデルを利用した平衡分布からのサンプリング

## III. 大きなスケールのシミュレーションの高速化

これらに続いて量子化学より大きなスケールのシミュレーションとして、分子動力学シミュレーション・有限要素法・流体シミュレーションのAIを使った高速化についても取上げる。

- 第9回 | 分子動力学シミュレーションの高速化
- 第10回 | 有限要素法に基づくシミュレーションの高速化
- 第11回 | 液体シミュレーションの高速化

これらは既存の物理知識やシミュレーション技術とAIをいかに融合させるのが課題となる。これらについても紹介していこう

## IV. AIと対称性

最後に全体を通して重要となる対称性とAIについて紹介したい。対称性は物理法則や保存則などとも密接に関係し科学の進展とも関係してきた。AIに対称性をいかに導入するか、逆にデータから対称性をいかに発見するかといった最新状況についても紹介したい。

- 第12回 | AIと対称性

これらのトピックを通じて、最近のAIと化学との接点、および現時点での限界についても明らかにする予定である。

それでは本連載を通じて基礎となる機械学習とニューラルネットワークについて紹介していく。

### 機械学習の基礎

ここから機械学習とニューラルネットワークについて紹介していくが、詳しく知りたい方は文献1, 2などを参考にしてほしい。

**機械学習**は学習データからルールや法則を帰納的に獲得するようなアプローチである。この実現には学習データ、モデル、そして目的関数の三つを用意する。学習データは入力と正解となる出力のペアからなる複数のデータである。モデルは（一般に）パラメータで特徴づけられた関数であり、パラメータを変えると関数の入出力の関係が変わる。そして、目的関数は学習データとモデルを受取って、モデルによる予測があたっていれば小さな値を返すような関数である。データごとに定義された損失関数（正解値と予測値のズレの大きさ）の平均として目的関数が定義される場合が多い。

そして学習とは目的関数が小さくなるようなモデルのパ

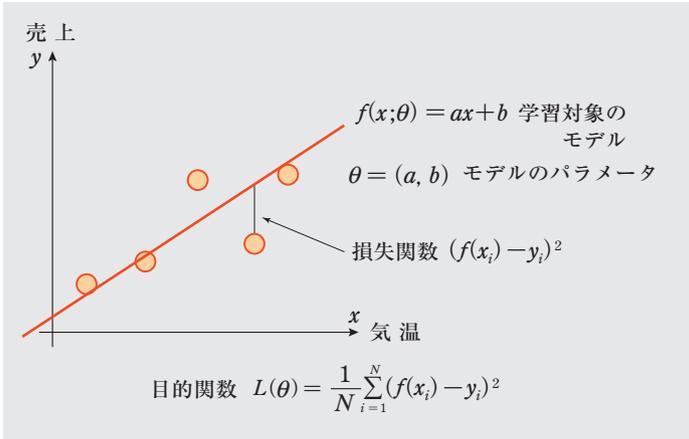


図1 学習データ（ある日の気温とその日の売上）から学習対象のモデルを学習する例。ここでは1次関数  $f(x) = ax + b$  を学習対象モデルとし、そのパラメータ  $\theta$  は  $a, b$  である。モデルによる予測と正解との二乗誤差を損失関数とし、その平均が目的関数である。目的関数を  $\theta$  について最小化することで学習結果のモデル  $f(x; \theta^*)$  を得る。

パラメータを求める最適化問題として定式化される。目的関数が小さいということは学習データの入力から出力をうまく予測できているモデルということになる。

たとえば、入力  $x$  をある日の気温、出力  $y$  をその日のアイスクリームの売上とし、気温から売上を予測するタスクを考える（図1）。学習データを  $N$  個の入力と出力のペア  $\{(x_i, y_i)\}_{i=1}^N$  とする。また、モデルは1次関数  $f(x; \theta) = ax + b$  とする。ここで  $\theta = (a, b)$  がモデルのパラメータである。目的関数  $L(\theta)$  はモデルによる予測と正解との二乗誤差を損失関数とした平均

$$L(\theta) = \frac{1}{N} \sum_{i=1}^N (f(x_i) - y_i)^2$$

とする。学習は、この目的関数を最小化するようなパラメータ  $\theta^* = (a^*, b^*)$  を求めることになる。

この例の場合は1次関数であるため、最適なパラメータ  $a^*, b^*$  は解析的に求まるが、一般にはモデルが非線形な形をしていたりデータが大量にあるため、計算量の問題から最適なパラメータが求まらない。そのため、学習における最適化手法の研究が進んでいる。たとえば、一部のデータを使って  $L(\theta)$  の  $\theta$  についての勾配の近似値を求め、それに従ってパラメータを更新する確率的勾配降下法を用いる。

本連載では機械学習を用いて、たとえば原子配置からエネルギーや力を推定したり、アミノ酸配列からタンパク質構造を推定したりする。また平衡分布からの生成を学習目

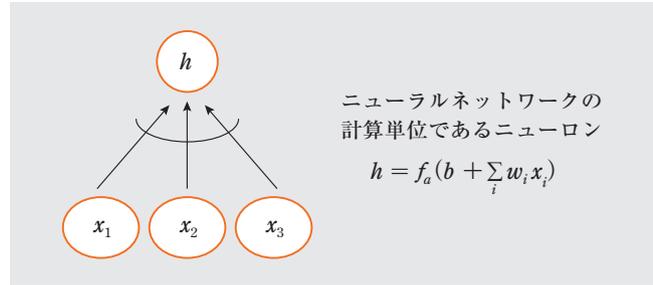


図2 ニューロンの計算。ニューロンの状態はほかの接続するニューロン  $x_i$  からの出力を重み  $w_i$  をかけたうえでバイアス  $b$  を加え、非線形の活性化関数  $f_a$  を適用した結果とする。

標とし、平衡分布からのサンプリングを実現する例なども見ていく。

## ニューラルネットワーク

現在の機械学習モデルにおいてはニューラルネットワークを用いることが多くなっている。ニューラルネットワークは単純な計算を行うニューロンとよばれる計算単位から構成されている（図2）。各ニューロンの状態（活性値）は、それが接続するほかのニューロンの状態を重みづけをしたうえで合計値を求めたあと、活性化関数とよばれる非線形関数を適用した結果で求まる。こうしたニューロンを横に並べた層を何層も並べてニューラルネットワークが構成される（図3）。入力が1層目、出力が最終層に対応する。ベクトルを入力とし、線形変換を施したあと、非線形関数で変換すると考えてもよい。モデルのパラメータは接

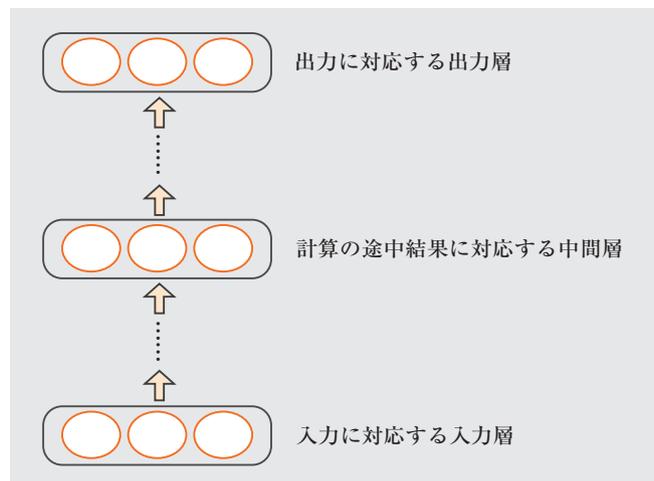


図3 ニューロンを横に並べた層を何層も重ねてニューラルネットワークが構成されている。入力に対応する入力層、出力に対応する出力層、計算の途中結果に対応する複数の層から構成される。

続の重みである。

ニューラルネットワークは単純だが、どれだけ複雑な入出力の関係がある関数であってもいくつかの条件を満たした十分大きなニューラルネットワークであれば、任意の精度でその関数を近似できることがわかっている（**万能近似定理**）。

さらに、真の関数の形状が、不連続な関数であったり、滑らかさが非一様であったり、データが低次元上に分布する場合は**カーネル法**や少ない層数のニューラルネットワークより多くの層からなるニューラルネットワークの方が有利であることがわかっている（文献3, 5）。

ニューラルネットワークは**誤差逆伝播法**とよばれる方法で、（出力の先にある）目的関数の各パラメータや状態についての勾配を正確かつ効率的に求めることができる。これを使って学習に必要な勾配情報が計算できる。

ニューラルネットワークではニューロン間をどのように接続するのか、またパラメータをどのように共有するのかによって学習性能や汎化性能が大きく変わる。たとえば化学分野ではグラフ構造をそのまま取込んだグラフニューラルネットワークや、Transformerとよばれるモデルが使わ

れることが多い。

また、本連載では対称性を多く扱う。これは座標のとり方を変えても法則は変わらないといった入力や出力に対する作用に対する不変性、同変性（共変性、反変性など）に関してもモデルの工夫で扱えることを見ていく。

## まとめ

第1回では計算化学とAIについての背景、連載の展望および機械学習とニューラルネットワークの基礎を紹介した。次回はニューラルネットワークポテンシャルを紹介する。

## 参考文献

1. 岡野原大輔 著, 『ディープラーニングを支える技術』, 技術評論社 (2022).
2. 岡野原大輔 著, 『ディープラーニングを支える技術2』, 技術評論社 (2022).
3. R. Nakada, M. Imaizumi, *J. Mach. Learn. Res.*, 21 (174) , 1 (2020).
4. T. Suzuki, A. Nitanda, “Deep learning is adaptive to intrinsic dimensionality of model smoothness in anisotropic Besov space”, *NeurIPS* 2021.
4. M. Imaizumi, K. Fukumizu, *PMLR*, 89, 869 (2019).