

幾何学的深層学習の分子・結晶解析への応用

谷 合 竜 典

1. はじめに

私たちの世界は、数えきれないほど多様な物質が、複雑に影響し合いながら存在することで成り立っています。これらの物質の正体は、様々な原子が結びつき、空間的に配置された「3次元構造」に他なりません。分子やタンパク質、鉱物や金属など、あらゆる物質の性質や機能は、この微小構造に支えられています。

物質科学の分野では、物質構造の性質や機能を解明・応用することで新しい薬やデバイス材料の開発が進められてきました。しかし、従来の手法は実験的な試行錯誤や偶然に頼る側面が大きく、多大な時間とコストを要します。これに対し、量子力学や統計力学に基づく第一原理計算や分子動力学法などの物理シミュレーションが登場し、物質の性質を理論的に予測する有力な手段となりました。しかし、依然として計算負荷や適用範囲などに課題があるほか、適切なモデル構築には、物性の発現原理に対する根本的理解が求められます。

こうした物質科学の分野にも、近年、深層学習技術の発展による新たな変革が訪れています。機械学習を用いることで、従来の物理シミュレーションよりもはるかに高速に物質の性質を予測できる¹⁻⁴⁾だけでなく、新たな物質構造の創出⁵⁻⁷⁾や、人間の直感では到達できないスケールでの広大な物質空間の探索^{8,9)}までもが現実になりつつあります。

こうした革新を支える基盤となるのが、物質構

造を「入力」し、性質や構造に対する作用を「出力」する表現学習において、構造の3次元的な幾何学性や回転・並進操作に対する対称性といった物理的性質を尊重する、幾何学的深層学習 (geometric deep learning)^{10,11)} というアプローチです。

本稿では、この分野の基礎となる考え方や回転群の対称性、代表的な機械学習手法の枠組みなどを紹介しながら、筆者らの最新の研究成果である Crystalformer³⁾ および CrystalFramer⁴⁾ に基づく結晶構造解析の応用例を解説していきます。

2. 原子から成る物質の3次元構造

物質を構成する最小単位である原子は、単独で存在することはまれで、通常は複数の原子が結びつくことで安定な構造を形成します。こうしてできた原子の集まりは、図1に示すように、分子やタンパク質、結晶といった多様な形態を取り、それぞれ異なる特徴を備えています。

本節では、機械学習で扱う際のデータ表現にも着目しながら、これらの構造の特徴を紹介します。



図1 代表的な物質構造の種類。

表 1 水 (H_2O) の分子構造.

原子種	x	y	z
H	0.000	-0.755	-0.471
O	0.000	0.000	0.118
H	0.000	0.755	-0.471
結合ペア	結合状態		
1-2	共有結合		
2-3	共有結合		

2.1 分子構造

化学物質の基本単位である分子は、数個から数百個程度の比較的少数の原子が、共有結合やイオン結合、金属結合といった化学結合によって結びついた構造です。分子内の原子の配置や、結合の種類・角度は分子固有の 3 次元形状を形づくり、その形が性質や機能に大きな影響を与えます。

たとえば、水分子 H_2O (図 1 左) は、中心の酸素原子に対して水素原子が約 104.5° の角度で配置された折れ曲がった構造をしており、これが水の様々な性質の起源となっています。

構成原子が同じでも空間配置が異なる「立体異性体」という概念もあり、中でも「鏡像異性体 (光学異性体)」は、互いに鏡像の関係にある構造であり、片方が無害でももう片方は有害といった場合も存在します。これらの構造の違いを捉えるには、回転や対称性の概念 (第 3 節) が重要になります。

分子構造は、原子の 3 次元的な配置を点群データとして表すことができます。表 1 では、 H_2O 分子を構成する 3 つの原子について、それぞれの 3 次元座標 (単位はオングストローム、 $\text{\AA} \equiv 0.1 \text{ nm}$) を示しています。さらに、原子リストの中で、どの原子同士が結合しているかを表す原子間の接続情報も、同表の下部に付記しています。

2.2 タンパク質構造

タンパク質は、アミノ酸が鎖状に連なった 1 次元構造を持ち、これが折りたたまれて複雑な 3 次元構造を形成します (図 1 中央)。タンパク質を構成するアミノ酸は一般に 20 種類あり、アミノ酸 1 つにはおよそ 10 から 30 個の原子が含まれます。これが数百個連なることで、全体として数百から数万個の原子からなる巨大な分子が構成されます。

タンパク質も分子の一種ではありますが、その構造的な複雑さや、生体内で担う機能の多様性・重要性から、一般的な低分子化合物とは区別して扱われることが多いです。

タンパク質分野では、アミノ酸配列から 3 次元構造を予測する問題は「タンパク質の折りたたみ問題」として長年にわたり研究されてきた難問でした。近年、DeepMind 社が開発した深層学習モデル AlphaFold により高精度な構造予測が可能となり、この分野は飛躍的な進展を遂げました。

2.3 結晶構造

鉱物や金属、半導体材料などは、原子が規則的に配列した 3 次元の周期構造、すなわち結晶構造を形成します。結晶構造は、**単位胞** (unit cell) と呼ばれる最小の構造パターンを 3 次元空間に繰り返すことで記述されます。単位胞は通常、平行六面体の形をしており、数個から数百個の原子を含みます。この単位胞の六面体の形状や内部の原子配置が、結晶全体の構造や物性を決定づけます。

結晶構造は高い対称性を持つことが知られており、**空間群** (space group) という概念で体系的に分類されています。たとえば、塩化ナトリウム (NaCl) 結晶の単位胞 (図 1 右) は、 Na^+ イオンと Cl^- イオンが交互に並ぶ立方格子構造を持ち、両イオンが対称的に配置されています。これは面心立方構造と呼ばれ、対称性が高い構造の一つです。

結晶構造データは、単位胞内の原子配置と、単位

表 2 塩化ナトリウム (NaCl) の結晶構造.

原子種	x	y	z
Na	0.000	0.000	0.000
Na	0.000	2.820	2.820
Na	2.820	0.000	2.820
Na	2.820	2.820	0.000
Cl	2.820	2.820	2.820
Cl	2.820	0.000	0.000
Cl	0.000	2.820	0.000
Cl	0.000	0.000	2.820
格子ベクトル	x	y	z
ℓ_1	5.640	0.000	0.000
ℓ_2	0.000	5.640	0.000
ℓ_3	0.000	0.000	5.640

胞の形状を表す3つの格子ベクトル (lattice vectors) $\ell_1, \ell_2, \ell_3 \in \mathbb{R}^3$ によって記述されます. 表2では, NaCl の単位胞を構成する原子の3次元座標 (上段) と対応する格子ベクトル (下段) を示しています. 原子数に関して, 図1 (右) では単位胞の境界上に位置する原子も含めて可視化していますが, データとしての単位胞 (表2) には, 原子の位置が $0 \leq x, y, z < 5.640$ (Å) の単位胞領域に収まる8個の原子のみが含まれます. また, 分子の場合とは異なり, 結晶構造では一般に原子間の結合を明示的に記述することはありません.

3. 回転群・並進群と対称性

3次元構造データを対象とする機械学習において, 回転や並進といった変換操作や対称性の概念は極めて重要です. これは, 表1や表2に示すように, 物質構造が3次元座標に基づく原子の点群データとして表現される一方で, その座標系の選び方には自由度 (任意性) があるためです.

入力データの座標系が異なっても, 機械学習モデルが本質的に同じ挙動を示す場合, これはモデルが特定の操作に対して**対称性** (symmetry) を有すると表現されます. モデルの振舞いに関して「対称性」という言葉を用いることに対し, 幾何学的な対称性のイメージとは異なる印象を持つ方もいるかもしれません. しかし, 幾何学における対称性も, 回転や反転といった操作を加えても図形の形が変わらないという点で, 本質的には同じ概念と捉えることができます.

本節では, 幾何学的深層学習において重要な役割を果たす, 3次元空間の回転群と並進群, そしてそれらの対称性の概念について紹介します.

3.1 回転群・並進群の数学的基礎

n 次元ユークリッド空間における回転操作 R は, 一般に $RR^T = I$ および $\det R = 1$ を満たす直交行列 $R \in \mathbb{R}^{n \times n}$ による座標変換 $\mathbf{p}' = R\mathbf{p}$ として記述されます. 回転操作全体の集合を $G(n)$ としたとき, 回転操作は次の4つの演算則を満たします.

1. **結合則**: 任意の順で回転を合成しても, 結果は変

表3 3次元の回転操作にまつわる群.

群	回転	鏡映	並進
$SO(3)$	含む	含まない	含まない
$O(3)$	含む	含む	含まない
$SE(3)$	含む	含まない	含む
$E(3)$	含む	含む	含む

わらない. すなわち, $(R_1 R_2) R_3 = R_1 (R_2 R_3)$.

2. **単位元の存在**: 単位操作 I (回転角0の回転) が存在し, $IR = RI = R$ を満たす.

3. **逆元の存在**: 各 R に対して逆操作 $R^{-1} = R^T$ が存在し, $R^{-1}R = R^T R = I$.

4. **閉包性**: $R_1, R_2 \in G(n)$ のとき $R_1 R_2 \in G(n)$.

以上より, 回転操作は群の公理を満たしており, これらの操作全体の集合は**回転群**と呼ばれます. 特に, 3次元空間における回転群は, $SO(3)$ 群 (special orthogonal group in 3D) と呼ばれます.

対象とする直交変換 R に, $\det R = -1$ を満たす鏡映変換 (反転操作) も含めた場合, この集合も群をなし, $O(3)$ 群 (orthogonal group in 3D) と呼ばれます. さらに, 回転に並進操作 $\mathbf{t} \in \mathbb{R}^3$ を加えた座標変換 $\mathbf{p}' = R\mathbf{p} + \mathbf{t}$ を考えると, 新たな群構造が得られます. 特に $SO(3)$ に対応するのが $SE(3)$ 群 (special Euclidean group in 3D), $O(3)$ に対応するのが $E(3)$ 群 (Euclidean group in 3D) です. これらの関係を表3にまとめます.

3.2 対称性: 不変性と同変性

先ほど, 機械学習モデルが座標系の表面的な違いにとらわれず, 本質的に同じ挙動を示すことを, 「対称性を有する」と表現しました. 物質構造を対象とする機械学習モデルを設計する際には, 先述の $SO(3)$ 群や $SE(3)$ 群といった変換操作に対して, このような対称性を備えることが望まれます.

ここで, 「本質的に同じ挙動」という表現には, 実は性質の異なる2種類の対称性の概念が含まれています. それが**不変性** (invariance) と**同変性** (equivariance) です.

たとえば, N 個の原子からなる物質構造の原子座標を, 行列 $P = [\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_N] \in \mathbb{R}^{3 \times N}$ で表すとします. ここで各列 \mathbf{p}_i は, 3次元空間における原子 i の位置ベクトルです. P を入力とする機

機械学習モデルを関数 f とした場合、 f が **SE(3) 不変性**を備えるとは、任意の回転行列 $R \in SO(3)$ および並進ベクトル $t \in \mathbb{R}^3$ に対して、次の関係が成り立つことを意味します。

$$f(P) = f(RP + [t \cdots t]). \quad (1)$$

このように、入力が一定の変換操作（ここでは回転と並進を含む **SE(3) 変換**）を受けたとしても、モデル f の出力が変わらないという性質が、不変性です。たとえば、構造の持つエネルギーなどのスカラー物性値の推定には、このような不変性を有するモデルが有効となります。

一方で、同変性は、ベクトルや構造変化などの幾何学的量を推定する際に重要になります。たとえば、構造 P に対して、より安定な構造 $P' = \Delta P + P$ を推定する問題を考えます。このとき、構造の変分 ΔP をモデル f により $\Delta P = f(P)$ と予測するとします。理想的には、入力構造 P を回転させた場合、その変換に応じて出力の変分 ΔP も同様に回転することが望まれます。すなわち、

$$Rf(P) = f(RP) \quad (2)$$

が成り立つことが望まれ、このような性質は回転操作に対する同変性、特に **SO(3) 同変性**と呼ばれます。より一般には、式 (2) は、回転 R に応じた作用 $g(R)$ を通じて $g(R)f(P) = f(RP)$ と記述され、不変性は同変性において $g(R) = 1$ とする特殊ケースと考えることができます。

3.3 機械学習における幾何学的対称性の重要性

対称性の概念は、幾何学的データに内在するパターンを効果的に学習する上で重要です。深層学習モデルでは、入力 x に対して $y = \max(w^T x + b, 0)$ のような線形変換と非線形変換を繰り返し適用することが基本となります。このとき y は、 x がある超平面の片側にあるときのみ活性化されます。このような超平面をいくつも用いることで、入力空間を細かく分割し、領域ごとの変換を学習します。

では、入力 x が点の座標 p だった場合はどうでしょうか。例えば、表 1 の水分子の 3 次元構造に回転や並進操作を施し、視覚的には図 2 の左から右の状態へ変換した場合を考えます。このように

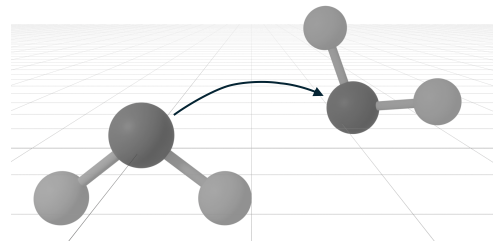


図 2 H₂O 分子構造の回転・並進の様子。

表 4 回転・並進した H₂O 分子構造。

原子種	x	y	z
H	-1.709	1.838	1.035
O	-1.530	2.149	0.148
H	-1.055	2.970	0.280

視覚的に表現すれば、人間には同じ構造だと直感的に認識できます。しかし、変換前後の構造の座標データ（表 1 および表 4）を見比べて、それらが同じ構造を表していると判断するのは困難でしょう。コンピュータにとっても同様に、このような座標データを入力とする場合、単純な処理ではまったく異なる入力として扱われてしまうのです。

モデルが幾何学的な対称性を備えるとは、こういった座標変換後も構造を同一視できることであり、物理的整合性の保証に加え、学習データ効率や汎化性能の向上につながるのです。

4. GNN による幾何学的深層学習

物質構造の 3 次元情報を扱うには、原子の位置と種類からなる点群に、近傍原子との接続情報を加えたグラフを構築し、これを入力とする**グラフニューラルネットワーク (GNN)**によって物性を予測する手法が一般的です^{1, 2, 11)}。最も基本的な分子向け GNN の模式図を、図 3 左に示します。

ここでは、GNN の目的を、シミュレーションなどで得られた物性値を教師ラベルとして学習し、それを高速に予測することとして説明を進めます。

GNN では、各原子の状態を抽象的な特徴ベクトル $x \in \mathbb{R}^d$ で表し、構造全体の状態を $X = [x_1, x_2, \dots, x_N] \in \mathbb{R}^{d \times N}$ として保持します。各 x は、入力時点では原子種 (H や O など) を記号的

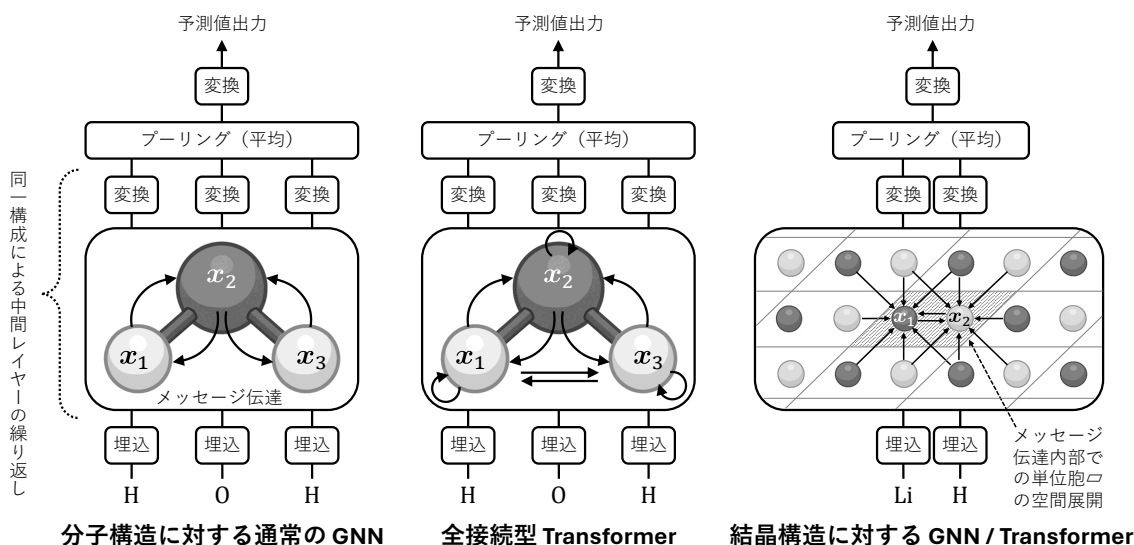


図 3 分子構造および結晶構造に対するニューラルネットワーク。

に表した埋め込みベクトルとして符号化され^{*1)}、ネットワーク内の様々な変換を通じて、空間内での原子配置に応じた状態へと更新されていきます。

中でも特に重要なのが、以下のような、**原子間相互作用** (interatomic interaction) をモデル化する**メッセージ伝達** (message passing) です。

$$\mathbf{x}'_i = \sum_{j \in N_i} w_{ij} \mathbf{m}_{i \leftarrow j}(\mathbf{x}_j, \mathbf{p}_j - \mathbf{p}_i). \quad (3)$$

ここで $j \in N_i$ は原子 i の近傍、 w_{ij} は重み係数、 $\mathbf{m}_{i \leftarrow j}$ は原子状態と相対位置に基づくメッセージ関数で、 j から i への作用を抽象的に表します。

通常、近傍原子集合 N_i は、原子間の結合関係や、距離に基づく k 近傍などによって定義されます (図 3 左)。一方、**Transformer**¹²⁾ に基づくモデル (図 3 中央) では、 N_i を i 自身を含む全原子とし、重み w_{ij} を状態変数 \mathbf{x} に基づいて動的に計算する**自己注意** (self-attention) 機構を用います。

このようなメッセージ伝達と、各原子独立の状態変換 (例: $\mathbf{x}' = \mathbf{x} + W_2 \max(W_1 \mathbf{x} + \mathbf{b}_1, \mathbf{0}) + \mathbf{b}_2$) を基本構成ブロックとして、GNN 内で繰り返し適用します。構造全体のエネルギーなどを予測す

る場合は、平均処理 $\sum_i \mathbf{x}_i / N$ などのブーリング処理を経て、最終的な予測値が出力されます。

幾何学的 GNN では、式 (3) のメッセージ伝達における座標データの対称性の扱いが重要になりますが、並進対称性は相対位置表現 $\mathbf{p}_j - \mathbf{p}_i$ を用いることで自然に実現されます。したがって、幾何学的 GNN では、回転対称性の扱いが主な課題となります。以下では、不変性や同変性を実現するための代表的な 3 つのアプローチを紹介します。

4.1 不変特微量

回転を含む $SE(3)$ 不変なモデルを実現する最も単純な方法は、不変な幾何学的量のみを用いてメッセージ伝達を設計することです。距離 $r_{ij} = |\mathbf{p}_j - \mathbf{p}_i|$ は最も代表的な不変量で、多くの手法が採用しています^{1~3)}。しかし、距離情報は 3 次元の情報も多くを損ないます。そこで、不変性を保ちながらモデルの表現力を向上する研究が多くされています。たとえば、メッセージ伝達において、 i とその近傍 $j, k \in N_i$ の間の三体関係をモデル化する方法では、 $i \rightarrow j$ と $i \rightarrow k$ の 2 つの方向間の角度を不変量として利用できます。しかし、三体作用を扱うモデルでは、 (i, j, k) の組合せ数が多くなり、計算量が肥大化する傾向があります。

4.2 フレーム

回転対称性は、入力構造 P の向きを予め正規化

*1) 原子埋め込みの最も簡単な方法としては、構造中の各原子 i に対して、原子番号に対応する次元のみが 1 となる one-hot ベクトル $\mathbf{a}_i = [\dots 0 \ 1 \ 0 \dots]^T$ に、学習可能な重み行列 W を掛ける処理 ($\mathbf{x}_i = W \mathbf{a}_i$) によって実現できます。

することでも実現できます．たとえば， P に対して主成分分析を適用すれば，正規直交な固有ベクトルからなる行列 $F = [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3]^T$ が得られ，構造を FP のように標準化できます．このように構造に追従して定まる座標系 F を**フレーム** (frame) と呼びます．ただし，固有ベクトルの符号には任意性があるため， F は一意には定まらず，鏡映変換を除いても通常 4 通りの候補が存在します．

一般に， P に対応するフレームの集合 $\mathcal{F}(P)$ が有限である場合，以下の**フレーム平均化** (frame averaging)¹³⁾ によって，任意のモデル f を不変化した \bar{f} を得られます．

$$\bar{f}(P, X) = \sum_{F \in \mathcal{F}(P)} \frac{1}{|\mathcal{F}(P)|} f(FP, X). \quad (4)$$

フレームは，不変量だけでなく，同変量の推定にも利用できます．たとえば， f の出力が変分 ΔP のような幾何学的量である場合，式 (4) の右辺において $f(FP, X)$ を $F^{-1}f(FP, X)$ に置き換え，フレーム逆変換によって元の座標系に戻すことで，同変性を実現できます．

4.3 同変特微量

上記の不変表現に基づく 2 つのアプローチとは対照的に，回転情報を保持したまま処理を行う**同変ネットワーク**の研究も進んでいます．これらのモデルでは，**球面調和関数** (spherical harmonics) の回転同変性に基づいた同変特微量を用います．

球面調和関数とは，単位球面上に定義された関数をフーリエ級数展開する際の直交基底です．通常の周期長 T の 1 次元信号 $f(t)$ に対するフーリエ級数展開では， $\omega_0 = \frac{2\pi}{T}$ を基本角周波数とした三角関数基底 $\{\exp(ik\omega_0 t)\}_{k \in \mathbb{Z}}$ の重ね合わせによって $f(t)$ を表現します．この変換を球面上の関数 $f(\bar{\mathbf{r}})$ に拡張した際に登場する直交基底関数 $\{Y_l^m(\bar{\mathbf{r}})\}$ が，球面調和関数 (図 4) です．ここで $\bar{\mathbf{r}} \in \mathbb{R}^3$ は 3 次元単位ベクトルで，次数 $l \in \{0, 1, 2, \dots\}$ は角周波数に相当し，各 l に対して $m \in \{-l, -l+1, \dots, l\}$ の $2l+1$ 個の基底関数が存在します (図 4 では縦方向が l ，横方向が m に対応)．球面調和関数を用いると，任意の滑らかな球面関数 $f(\bar{\mathbf{r}})$ を，各 Y_l^m

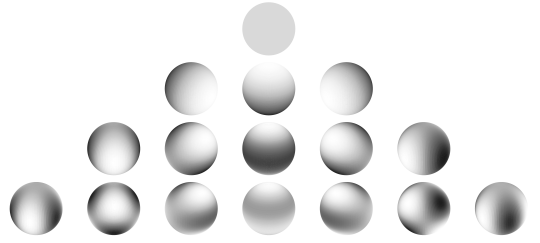


図 4 球面調和関数 (白が正，黒が負の値領域)．

に対応する係数 x_l^m により表現できます^{*2)}．

$$f(\bar{\mathbf{r}}; \mathbf{x}) = \sum_{l=0}^{\infty} \sum_{m=-l}^l x_l^m Y_l^m(\bar{\mathbf{r}}). \quad (5)$$

1 次元信号 $f(t)$ のフーリエ成分を $F(\omega_k)$ とすると，並進操作 $f(t - t_0)$ によって各成分は $\exp(-i\omega_k t_0)F(\omega_k)$ となります．これはフーリエ表現が並進同変であることを示します．同様に，球面に沿った座標シフト，すなわち回転操作 R に対して，球面調和関数は次の同変性を持ちます．

$$D^{(l)}(R)Y_l(\bar{\mathbf{r}}) = Y_l(R\bar{\mathbf{r}}). \quad (6)$$

ここで， $Y_l(\bar{\mathbf{r}})$ は次数 l における球面調和関数の成分を並べた $2l+1$ 次元ベクトル， $D^{(l)}(R)$ は回転 R に対応する**ウィグナーの D 行列**と呼ばれる $2l+1$ 次元のユニタリ行列です．

式 (5) および (6) を合わせて考えると，球面調和関数基底の係数 $\{x_l^m\}$ によって，球面全体にわたる関数応答を表現できるだけでなく，入力構造の回転状態も同時に符号化できることがわかります．

同変ネットワークではこの性質を利用し，各原子の状態として，次数 $l = 0, 1, \dots, L$ とその成分 m を持つピラミッド構造の同変特微量 $\bar{\mathbf{x}} = \{\mathbf{x}_l^m\}$ を保持します．そして，全体の特微量 $\bar{\mathbf{X}}$ に対して回転同変性 $D(R)h(\bar{\mathbf{X}}) = h(D(R)\bar{\mathbf{X}})$ を満たす変換レイヤー h を通じて，特微量 $\bar{\mathbf{X}}$ を更新していきます．特に式 (3) のメッセージ伝達では，相対位置 $\mathbf{r}_{ij} = \mathbf{p}_j - \mathbf{p}_i$ を不変な距離成分 r_{ij} と，同変な球面調和関数成分 $\mathbf{Y}(\bar{\mathbf{r}}_{ij})$ に分解し，これら構造情報を**テンソル積**などの演算を介してメッセージ

*2) $Y_l^m(\bar{\mathbf{r}})$ や x_l^m は通常，複素数値 (C) を取りますが，GNN に実装する際には実数値版の Y_l^m (図 4) が用いられます．

関数 $\mathbf{m}_{i \leftarrow j}$ に組み込みます。最終的に出力層において $l = 0$ の成分を用いることで不変量, $l = 1$ の成分で 3 次元ベクトル等の同変量を推定します。

同変ネットワークには、同変性を保つために非線形変換の形式に強い制約があるほか、次数上限 L を上げると計算コストが急増するなどの限界もあります。同変ネットワークについての詳細な解説は、文献 [10, 11] などを参照してください。

5. 結晶構造の物性予測への応用

幾何学的深層学習の応用として、結晶構造の物性予測問題を考えます。結晶構造は、単位胞内の原子の種類 $A = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_N]$, 位置 $P = [\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_N]$, そして格子ベクトル $L = [\ell_1, \ell_2, \ell_3]$ により, (A, P, L) と表されます。結晶内の原子は、単位胞内原子を並進した位置

$$\mathbf{p}_i(\mathbf{n}) = \mathbf{p}_i + n_1 \ell_1 + n_2 \ell_2 + n_3 \ell_3 \quad (7)$$

に配置され、この原子 $i(\mathbf{n})$ は単位胞内原子 i と同じ状態を持ちます。ここで、 $\mathbf{n} = (n_1, n_2, n_3)^T \in \mathbb{Z}^3$ は格子並進を表す整数ベクトルです。

エネルギーなどの物性は座標変換に対して不変なので、ここで考えるモデル $f(A, P, L)$ も $SE(3)$ 不変なものを考えます。第 4 節で述べたように、モデル f は、内部で各原子の抽象的な状態ベクトル \mathbf{x}_i を保持・更新することで所定の物性予測を行うよう設計され、教師あり学習で訓練されます。

5.1 周期構造を捉える幾何学的 Transformer

結晶構造の物性予測手法として Crystalformer³⁾ を紹介します。これは Transformer モデル¹²⁾ を標準的形式のまま結晶構造に適用した初めての例です。Transformer は元々は自然言語処理のために開発されたモデルですが、単語の意味が前後の文脈によって決まるという点は、原子が空間配置に応じて多様な状態を取る今回の問題と類似します。

文章を表す単語列に文脈を反映するのと同様に、結晶中の周囲の原子 $j(\mathbf{n})$ の配置や状態に応じて各原子 i の状態を更新する処理も、次の自己注意操作に基づくメッセージ伝達で実現できます。

$$\begin{aligned} \mathbf{x}'_i &= \sum_{j=1}^N \sum_{\mathbf{n} \in \mathbb{Z}^3} w_{ij}(\mathbf{n}) \mathbf{m}_{i \leftarrow j}(\mathbf{n}), \\ w_{ij}(\mathbf{n}) &= \frac{1}{Z_i} \exp \left(\frac{\mathbf{q}_i^T \mathbf{k}_j}{\sqrt{d_K}} + \phi_{ij}(\mathbf{n}) \right), \\ \mathbf{m}_{ij}(\mathbf{n}) &= \mathbf{v}_j + \psi_{ij}(\mathbf{n}). \end{aligned} \quad (8)$$

ここで、 $\mathbf{q}, \mathbf{k}, \mathbf{v}$ は入力 \mathbf{x} の線形射影、 d_K は \mathbf{k} の次元数、 Z_i は重みの和 $\sum_j \sum_{\mathbf{n}} w_{ij}(\mathbf{n})$ を 1 にする正規化項 (softmax 関数の正規化項) です。

結晶構造に対するメッセージ伝達 (図 3 右) では、構造全体の状態を単位胞内の有限個の原子の状態集合 X で表し、単位胞内原子 j と同じ状態 \mathbf{x}_j を持つ原子 $j(\mathbf{n})$ が、式 (7) に従って空間内に無限に繰り返されると考えます。このとき原子 i と周辺原子 $j(\mathbf{n})$ の位置関係は、相対位置 $\mathbf{r}_{ij}(\mathbf{n}) = \mathbf{p}_j(\mathbf{n}) - \mathbf{p}_i$ に基づく埋め込み $\phi_{ij}(\mathbf{n})$ および $\psi_{ij}(\mathbf{n})$ によって表現されます。特に、距離 $|\mathbf{r}_{ij}(\mathbf{n})|$ を用いることで、 $E(3)$ 不変性を保証することができます。

さらに式 (8) の収束を保証するために、重み $w_{ij}(\mathbf{n})$ 内の位置表現に $\phi_{ij}(\mathbf{n}) = -|\mathbf{r}_{ij}(\mathbf{n})|^2 / 2\sigma_i^2$ の形を採用し、重みが距離に応じて自然に減衰するように設計されています。減衰半径 σ_i は状態ベクトル \mathbf{x}_i から動的に推定され、無限級数 $\sum_{\mathbf{n}}$ は σ_i の大きさに応じて適切な範囲で計算されます。

この原子間相互作用モデルは、距離減衰する重み $w_{ij}(\mathbf{n})$ と作用 $\mathbf{m}_{i \leftarrow j}(\mathbf{n})$ の両方を高次元の状態変数上で定義することで、物理シミュレーションにおけるポテンシャル和の計算を抽象的に模倣しています。用いられるポテンシャル関数は単純なガウス型の距離減衰関数ですが、レイヤーを重ねることでより複雑な相互作用を柔軟に表現します。この手法は、多くの物性予測ベンチマークにおいて、既存手法の半分以下の学習パラメータ数でありながら、同等以上の予測性能を実現しました。

5.2 動的フレームに基づく構造の局所正規化

次に紹介する CrystalFramer⁴⁾ では、Crystalformer におけるメッセージ関数内の位置表現 $\psi_{ij}(\mathbf{n})$ に着目し、回転不変性を保ちながら原子間距離よりも豊かな幾何学的情報を取り込む手法として、「動的フレーム」を提案しています。

第 4.2 節で紹介したフレーム¹³⁾を用いれば、構造 P に対してフレーム変換 F を適用し FP とすることで、その向きを正規化し、回転不変な位置表現を得ることが可能です。しかし、結晶はもともと無限に広がる周期構造であり、また、単位胞 (P, L) の切り出し方にも任意性があるため、単位胞に基づいた安易な正規化には疑問が残ります。

この研究では、フレームの役割を単なる構造の正規化手段としてではなく、式 (8) におけるメッセージ関数 $\mathbf{m}_{i \leftarrow j(\mathbf{n})}$ の中で、相対位置 $\mathbf{r}_{ij(\mathbf{n})} = \mathbf{p}_{j(\mathbf{n})} - \mathbf{p}_i$ を効果的に取り込むための手段として再考しました。そこで、構造全体を一括で正規化する従来のフレームの代わりに、メッセージ伝達レイヤーごと・原子 i ごとに専用の動的フレーム F_i を構築します。このとき、原子 i から見て影響の小さい原子 $j(\mathbf{n})$ (すなわち重み $w_{ij(\mathbf{n})}$ が小さいもの) は、構造的に重要でないと見なせます。そこで、 $w_{ij(\mathbf{n})}$ を構造に対する“マスク”として用いて原子 i 周辺の局所構造を特定し、その領域の向きを正規化するようなフレーム F_i を構築します。

CrystalFramer では、Crystalformer で用いていた距離ベースの位置表現に加え、単位方向ベクトル $\bar{\mathbf{r}}_{ij(\mathbf{n})} = \mathbf{r}_{ij(\mathbf{n})}/|\mathbf{r}_{ij(\mathbf{n})}|$ をフレーム変換した $F_i \bar{\mathbf{r}}_{ij(\mathbf{n})}$ の各成分も、不変な位置表現として取り入れています。このようなシンプルな拡張のみで、モデルの表現力を大幅に向上させ、物性予測誤差を最大 14% 低減させることに成功しました。

6. おわりに

本稿では、物質の 3 次元構造を対象とする幾何学的深層学習について、回転や並進といった幾何学的操作と、それらに対する対称性の概念、そしてグラフニューラルネットワークによる機械学習アプローチを中心に解説しました。特に、結晶構造の周期性や回転・並進不変性に着目した幾何学的 Transformer モデル——Crystalformer³⁾ および CrystalFramer⁴⁾ の 2 手法を紹介し、分子より複雑な結晶構造に対して、その幾何学的情報をどのようにモデルに組み込むかを具体的に解説しま

した。

幾何学的深層学習は、物理的制約とデータ駆動の柔軟性を両立する手法として、物質科学や関連分野において今後も重要な役割を果たしていくと考えられます。特に構造生成モデルによる新物質発見⁵⁾ や人間の知識と構造情報の融合⁹⁾ は、新たな知見や技術の創出がますます期待されます。

参考文献

- 1) K. Schütt, P.J. Kindermans, H.E. Sauceda *et al.* SchNet: A continuous-filter convolutional neural network for modeling quantum interactions. In *NIPS 2017*.
- 2) T. Xie and J.C. Grossman. Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties. *Phys. Rev. Lett.* **120**, 145301 (2018).
- 3) T. Taniai, R. Igarashi, Y. Suzuki *et al.* Crystalformer: Infinitely Connected Attention for Periodic Structure Encoding. In *ICLR 2024*.
- 4) Y. Ito, T. Taniai, R. Igarashi *et al.* Rethinking the role of frames for SE(3)-invariant crystal structure modeling. In *ICLR 2025*.
- 5) C. Zeni, R. Pinsler, D. Zügner *et al.* A generative model for inorganic materials design. *Nature* **639**, 624–632 (2025).
- 6) R. Jiao, W. Huang, P. Lin *et al.* Crystal Structure Prediction by Joint Equivariant Diffusion. In *NeurIPS 2023*.
- 7) N. Chiba, Y. Suzuki, T. Taniai *et al.* Neural structure fields with application to crystal structure autoencoders. *Commun. Mater.* **4**, 106 (2023).
- 8) Y. Suzuki, T. Taniai, K. Saito *et al.* Self-supervised learning of materials concepts from crystal structures via deep neural networks. *Mach. Learn.: Sci. Technol.* **3**, 045034 (2022).
- 9) Y. Suzuki, T. Taniai, R. Igarashi *et al.* Bridging Text and Crystal Structures: Literature-driven Contrastive Learning for Materials Science. *Mach. Learn.: Sci. Technol.* (2025).
- 10) A. Duval, S.V. Mathis, C.K. Joshi *et al.* A Hitchhiker's Guide to Geometric GNNs for 3D Atomic Systems. arXiv: 2312.07511 (2024).
- 11) J. Han, J. Cen, L. Wu *et al.* A survey of Geometric Graph Neural Networks: Data Structures, Models and Applications. *Front. Comput. Sci.* (2025).
- 12) A. Vaswani, N. Shazeer, N. Parmar *et al.* Attention Is All You Need. In *NIPS 2017*.
- 13) O. Puny, M. Atzmon, E.J. Smith *et al.* Frame Averaging for Invariant and Equivariant Network Design. In *ICLR 2022*.

(たにあい・たつのり、オムロン サイニックス株式会社)