

平成 30 年 4 月 19 日現在

機関番号：14301

研究種目：基盤研究(B) (一般)

研究期間：2015～2017

課題番号：15H04116

研究課題名(和文) 網羅的第一原理計算とスパース推定による高精度原子間ポテンシャルの開発

研究課題名(英文) Procedure for constructing machine learning interatomic potentials

研究代表者

世古 敦人 (Seko, Atsuto)

京都大学・工学研究科・准教授

研究者番号：10452319

交付決定額(研究期間全体)：(直接経費) 12,500,000円

研究成果の概要(和文)：第一原理計算と機械学習手法に基づき、高精度原子間ポテンシャルを構築する手法の提案を行った。具体的には、系統的な構造記述子を導入することにより、遷移金属を含むすべての金属に対して高精度な原子間ポテンシャルを構築することが可能な統一的な枠組みを提案した。さらに、機械学習ポテンシャルと金属系で広く用いられているEAMポテンシャルとの関係を明らかにした。これは、物理的解釈が難しかった機械学習ポテンシャルの一つの解釈である。また、結晶粒界などの大規模構造への応用を目指し、構造データ分布に基づいた異常検知の機械学習による指標を導入することができた。

研究成果の概要(英文)：Machine-learning interatomic potential (MLIP) has been of growing interest as a useful method to describe the energetics of systems of interest. In the present study, we examine the accuracy of linearized pairwise MLIPs and angular-dependent MLIPs for 31 elemental metals. Using all of the optimal MLIPs for 31 elemental metals, we show the robustness of the linearized frameworks, the general trend of the predictive power of MLIPs and the limitation of pairwise MLIPs. As a result, we obtain accurate MLIPs for all 31 elements using the same linearized framework. This indicates that the use of numerous descriptors is the most important practical feature for constructing MLIPs with high accuracy. An accurate MLIP can be constructed using only pairwise descriptors for most non-transition metals, whereas it is very important to consider angular-dependent descriptors when expressing interatomic interactions of transition metals.

研究分野：計算材料科学

キーワード：機械学習 原子間ポテンシャル 分子動力学 回帰分析 第一原理計算

1. 研究開始当初の背景

第一原理熱力学計算

材料開発において、熱力学量・物性・相平衡関係などを、温度・圧力・化学組成の関数として、把握することは極めて重要である。近年、第一原理計算と統計熱力学的手法による有限温度計算を組み合わせた第一原理熱力学計算により、自由エネルギーや相平衡の計算が幅広く行われている。代表者はこれまでに網羅的第一原理計算、クラスター展開法、モンテカルロ計算を組み合わせた第一原理熱力学計算の技術開発を進めてきている。また、近年においては、第一原理計算と調和近似を組み合わせることにより、格子振動物性の計算が可能である。その結果、単純な結晶構造を持つ合金系に限らず、結晶構造が複雑な実用的に興味ある材料系において、熱力学物性や平衡状態図を高精度に計算できるようになった。しかし、多くの場合において、調和近似やクラスター展開の格子点近似を超えた取り扱いが必要となる。例えば、非調和格子振動自由エネルギーなどの熱力学量やイオン伝導体などにおける有限温度下における原子分布を計算する場合などが当てはまる。また、拡散や熱伝導率などの動的な物性を計算する場合も同様である。これらに対する一つの解決策は、統計熱力学計算の一つである分子動力学 (Molecular Dynamics, MD) 計算を用いることである。

分子動力学法における原子間力計算の現状

MD 法は、原子間力の計算をもとに運動方程式を解き、原子の集団的運動を計算する方法である。そのため、MD 法の計算精度は、主に原子間力の計算精度により決定される。最も高精度な原子間力の計算方法は、第一原理計算から直接原子に働く力を求め MD 計算を行う第一原理 MD 計算である。しかし、計算コストは非常に高く、例えばイオン伝導体などにおける有限温度下の原子分布を求めるような場合には、原子の運動が高速である高温域に限られる上、かなりのシミュレーション時間が必要となる。さらに計算できる原子数も百原子程度が限界であるなど、第一原理 MD 計算を適用できる系や現象は非常に限定的である。

第一原理 MD 計算を適用できない系や現象を対象とする場合は、原子間力の計算が非常に高速である原子間ポテンシャルを用いることが一般的である。しかし、物理的考察をもとにした原子間ポテンシャルの表式は単純であるため、原子間力の計算精度やポテンシャルの移植性が不足する場合が多い。このような理由から、計算精度と計算速度を両立させる原子間ポテンシャルを構築する方法が最近提案された。これは、数千から数万の規則構

造に対して第一原理計算による網羅的エネルギー計算を行い、そのデータからニューラルネットワーク法などの回帰分析手法により、結晶構造とエネルギーの関係式を推定するというものである。回帰分析手法とは、目的変数と説明変数の関係を統計的に推定する方法である。この枠組みにより、化学結合の種類にかかわらず、第一原理計算との誤差が 1 meV/atom 程度の原子間ポテンシャルを得ることができると期待される。

2. 研究の目的

第一原理計算と回帰分析に基づいた手法においては、結晶構造の情報を別の表現に変換したうえで、原子間ポテンシャルの構築が行われる。そのため、予測精度は結晶構造を表現する方法に大きく依存する。これは、結晶構造表現の最適化が重要であることを意味しているが、結晶構造表現に対する事前情報はなく、経験や試行錯誤により選ばれているというのが実情であった。代表者は、結晶構造表現の効率的最適化を行うため、Least Absolute Shrinkage and Selection Operator (LASSO) 回帰と呼ばれるスパース推定に基づいた高精度原子間ポテンシャルを提案した。この方法では、結晶構造を膨大な数の網羅的基底関数により表現し、基底関数とエネルギーを線形モデルによりつなげる。さらにスパース推定により、膨大な数の基底関数から、必要な基底関数だけを自動的に選択する。その結果、結晶構造表現を予備知識なく選び出すことができ、計算精度と計算速度を併せ持つ原子間ポテンシャルを構築することができる。本研究では、この網羅的第一原理計算とスパース推定に基づいた高精度原子間ポテンシャル構築技術の確立を目指す。

3. 研究の方法

まず、数千から数万程度のモデル構造について、第一原理計算を行い、モデル構造とエネルギー・原子に働く力・応力のセット (学習データ) を準備する。原子間ポテンシャルの予測能力向上のため、様々な体積や原子座標のモデル構造について第一原理計算を行う予定である。また、有限温度物性の計算精度を向上させるため、小さいセルかつ短時間の第一原理 MD 計算を行い、スナップショットのデータを学習データに含める。

次に、得られた学習データから、原子間ポテンシャルを構築する。その際、原子座標を数万程度の基底関数の組み合わせへと変換する。さらに、基底関数とエネルギーの関係を表す線形モデルを用い、基底関数とエネルギーの関係を LASSO 回帰により推定する。LASSO 回帰では線形の範囲内で、スパースな解 (多くの回帰係数がゼロの解) が得られるため、必要な基底関数のみ自動的に選択される。

最後に、構築した原子間ポテンシャルの予測精度や移植性を評価する。予測精度や移植性

を保証することは、原子間ポテンシャルを統計熱力学計算などへ応用する際に重要である。具体的には、学習データに含まれない多くの構造について第一原理計算を行い、これらのエネルギーと原子間ポテンシャルによる予測との差を評価する。また、凝集エネルギー、体積弾性率などの基本的な物性、構造相転移やフォノン分散関係などを評価し、第一原理計算から直接得られるものとの比較を行う。さらに、格子欠陥形成エネルギーなどを評価し、第一原理計算によるものと比較することにより、原子間ポテンシャルの移植性を評価する。予測精度が低い場合には、予測精度の低い構造を学習データに追加し、再度原子間ポテンシャルの構築を行うことにより、予測精度を改善する。

原子間ポテンシャル構築に適した新たな基底関数の考察を行う。原子間ポテンシャルの計算精度は基底関数に大きく依存するため、基底関数の考察は本研究の中でも重要な箇所である。すでに実施した単体 Na, Mg の原子間ポテンシャル構築では、原子間距離のみから導出されるコサイン型やガウス型の基底関数を用いた。コサイン型やガウス型以外の基底関数を導入することにより、多種多様な基底関数を用いた柔軟性の高いポテンシャル構築を目指す。多項式型、指数関数型、メキシカンハット型、モルレー型などを導入する。また、新たに導入した基底関数およびそれらに対応したエネルギーや原子間力、応力算出のためのプログラム開発を行う。その際、これまでに実施してきた第一原理熱力学計算におけるプログラミング経験を生かし、効率的なプログラム開発を行う。また、効率的にプログラム開発を行うため、無償で公開されているソースコードを積極的に利用する。

4. 研究成果

まず、Al, Cu, Zn など 10 種類の単体金属に適用することにより、第一原理計算と Lasso 原子間ポテンシャル構築手法の有効性および汎用性を検証した。それぞれの単体金属に対して BCC, FCC, HCP 構造などを変形させた 2700 構造について、平面波基底 PAW 法(VASP コード)による第一原理計算を行い、エネルギー、力、応力を評価した。次に、Lasso により、エネルギー、力、応力およびそれらに関する記述子のデータから回帰係数を推定した。構築した原子間ポテンシャルの予測能力を評価するため、新たに 300 構造の第一原理計算を行い、原子間ポテンシャルの誤差評価を行った。

図 1(a)に 10 種類の単体金属に対する Lasso ポテンシャルの予測誤差を示す。記述子およびそのべき乗に対応する 4836 個の回帰係数のうち、非ゼロ係数の個数を横軸に示しており、正則化パラメータ λ の値により調整される。前述のように、 λ を変化させることでエネルギー計算コストおよび予測精度の異なるポ

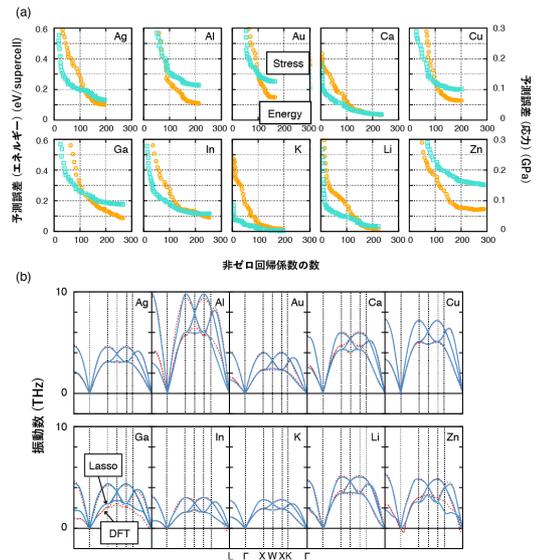


図 1 (a) Lasso ポテンシャルの予測誤差の非ゼロ回帰係数の数依存性. (b) Lasso ポテンシャルによるフォノン分散関係の予測. 実線が Lasso ポテンシャルによるもの、破線が第一原理計算によるものを示す。

テンシャルが得られる。それぞれの単体において、予測誤差が最小のポテンシャルは、100 個から 200 個程度の最適化された記述子で構成されている。構築した原子間ポテンシャルのエネルギー予測誤差は 1.0-3.5 meV/atom 程度、力や応力に対する予測誤差はそれぞれ 0.002-0.032eV/Å, 0.04-0.25 GPa 程度となった。ここでは、最適なものとして、エネルギー予測誤差が最小となるものを選んだが、エネルギーの予測精度と計算コストのバランスを考慮して、ポテンシャルを選ぶことが可能である。また、動力学特性の予測精度を示す一例として、図 1(b)に原子間ポテンシャルによるフォノン分散関係を示す。すべての単体において、フォノン分散関係は第一原理計算のものとはほぼ一致した。これらの結果は、第一原理計算と Lasso による原子間ポテンシャルにより、幅広い系において、エネルギーだけでなく原子に働く力や応力も高精度に予測できることを示している。

さらに、系統的な構造記述子を導入することにより、遷移金属を含むすべての金属に対して高精度な原子間ポテンシャルを構築することが可能な統一的な枠組みを提案した。さら

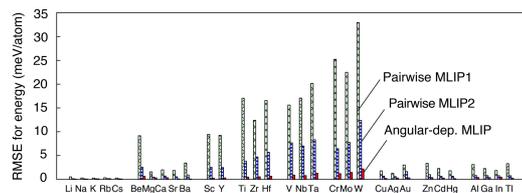


図 2 単体金属 31 元素における MLIP の予測誤差。

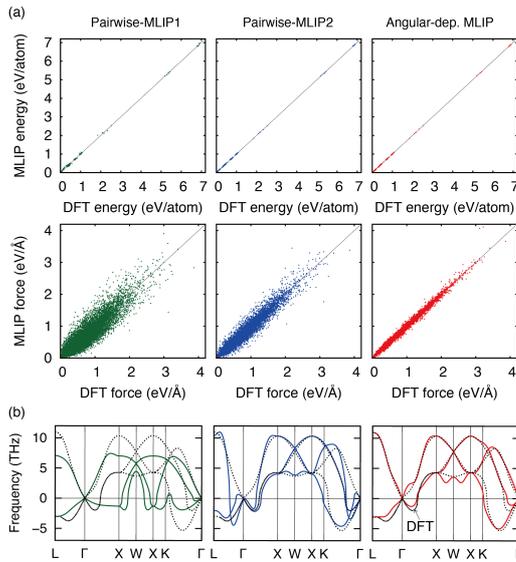


図3 Crにおけるエネルギー、力、フォノンに対する予測誤差。

に、機械学習ポテンシャル(MLIP)と金属系で広く用いられている EAM ポテンシャルとの関係を明らかにした。これは、物理的解釈が難しかった機械学習ポテンシャルの一つの解釈である。また、結晶粒界などの大規模構造への応用を目指し、構造データ分布に基づいた異常検知の機械学習による指標を導入することができた。

線形リッジ回帰法により、原子間距離のみの構造特徴量を用いた二種類の MLIP (Pairwise MLIP)と三体間角度の構造特徴量を加えた MLIP (Angular-dep. MLIP)を構築した。また、テストデータ (300 構造) について、エネルギーや原子に働く力に対する予測能力を評価した。さらに弾性定数、フォノン分散に対する MLIP の予測能力を評価した。

図2にてテストデータに対する MLIP の予測誤差を示す。Pairwise MLIP の典型金属元素に対する平均予測誤差は 0.7 meV/atom であり、Be や Hg を除く典型金属元素では、高精度な Pairwise MLIP を構築できた。一方で、遷移金属元素や Be, Hg では、三体間特徴量が不可欠である。一例として、図3にCrにおける予測誤差を示す。三体間特徴量の導入することにより、エネルギー、力、フォノンなどの予測精度が向上している事がわかる。最終的に、三体間特徴量を考慮した場合の平均予測誤差は 0.9 meV/atom となった。また、図4に示すように、弾性定数などの力学的特性も高精度に再現する MLIP を構築することができた。この結果は、すべての金属元素において、統一した枠組みにより高精度な MLIP を構築することができることを示している。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計3件)

- ① A. Seko, A. Takahashi and I. Tanaka, First-

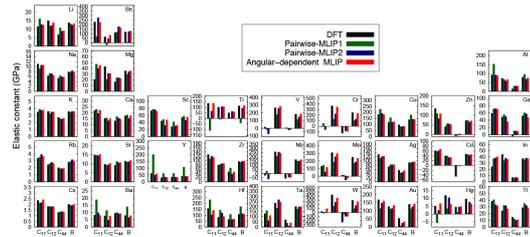


図4 FCC金属の弾性定数。

principles interatomic potentials for ten elemental metals via compressed sensing, Phys. Rev. B 92, 054113 (2015). 査読有

- ② A. Takahashi, A. Seko and I. Tanaka, Conceptual and practical bases for the high accuracy of machine learning interatomic potential, Phys. Rev. Materials 1, 063801 (2017). 査読有

- ③ A. Seko, H. Hayashi, K. Nakayama, A. Takahashi and I. Tanaka, Representation of compounds for machine-learning prediction of physical properties, Phys. Rev. B 95, 144110 (2017). 査読有

[学会発表] (計22件)

- (招待講演) A. Seko, Applications of machine learning to materials data, IPAM workshop: Optimization and Optimal Control for Complex Energy and Property Landscapes (2017).
- (招待講演) A. Seko, Linearized machine learning interatomic potentials for metals and recommender system for the discovery of unknown materials, Total Energy and Force Methods Workshop 2018 (2018).
- (招待講演) 世古敦人, 材料科学データに対する機械学習の応用, 加工プロセスによる材料新機能発現 第176委員会 第33回研究会 (2017).
- (招待講演) 世古敦人, 材料科学データにおける機械学習の応用, 日本物理学会共催シンポジウム (2018).
- (招待講演) 世古敦人, 材料科学データにおける機械学習の応用, 触媒学会コンピュータの利用研究会セミナー (2017).
- (招待講演) 世古敦人, 材料科学データに対する機械学習の応用, 情報・データ科学との連携・融合による物性物理・量子化学の新展開 (2017).
- (招待講演) 世古敦人, 材料科学データに対する機械学習の応用, 金属学会シンポジウム「データサイエンスと材料研究開発」 (2017).
- (招待講演) 世古敦人, 第一原理計算と機械学習手法を使った材料物性予測, 第26回格子欠陥フォーラム (2016).
- (招待講演) A. Seko, Applications of Machine Learning Techniques to First-principles Data, ENGE 2016 (2016).
- (招待講演) A. Seko, First-principles interatomic potentials via compressed sensing,

- TMS 2016 (2016).
11. (招待講演) A. Seko, First-principles interatomic potentials via compressed sensing, The 18th Asian Workshop on First-Principles Electronic Structure Calculations (2015).
 12. (招待講演) 世古敦人, 第一原理計算と圧縮センシングによる高精度原子間ポテンシャルの構築, 第6回計算統計物理学研究会 (2015).
 13. 世古敦人, 高橋 亮, 田中 功, 単体金属の線形回帰ポテンシャルに対する予測精度の一般的傾向, 日本金属学会 2018年春季講演大会 (2018).
 14. 近藤大介, 世古敦人, 田中 功, 大規模構造に対する線形回帰ポテンシャルの高精度化, 日本金属学会 2018年春季講演大会 (2018).
 15. 高橋亮, 世古敦人, 田中功, 線形回帰を用いた Ti の高精度原子間ポテンシャル構築, 日本金属学会 2017年秋期講演大会 (2017).
 16. 大槻祐介, 高橋亮, 世古敦人, 田中功, 精度保証付き線形回帰原子間ポテンシャルの構築について, 日本金属学会 2017年秋期講演大会 (2017).
 17. 高橋亮, 世古敦人, 田中功, 角度項を用いた線形回帰ポテンシャルの構築, 日本金属学会 2017年春季講演大会 (2017).
 18. 高橋亮, 世古敦人, 田中功, 線形回帰ポテンシャルによる単体金属の物性値予測, 日本金属学会 2016年秋季講演大会 (2016).
 19. 高橋亮, 世古敦人, 田中功, 単体金属における原子間ポテンシャル関数の線形表現, 日本金属学会 (2016)
 20. A. Seko, First-principles interatomic potentials via compressed sensing, MRS Fall meeting 2015 (2015).
 21. 高橋亮, 世古敦人, 田中功, 第一原理計算の線形回帰による原子間ポテンシャル構築:多体間相互作用の寄与, 日本金属学会 (2015).
 22. 世古敦人, 高橋亮, 田中功, 第一原理計算と圧縮センシングによる高精度原子間ポテンシャルの構築, 日本金属学会 (2015).

[図書] (計0件)

[産業財産権]

○出願状況 (計0件)

○取得状況 (計0件)

[その他]

ホームページ等

6. 研究組織

(1)研究代表者

世古敦人 (SEKO, Atsuto)

京都大学・工学研究科・准教授

研究者番号: 10452319