

2019年4月19日

報道関係者各位

高エネルギー加速器研究機構
東京理科大学
情報・システム研究機構 統計数理研究所

機械学習によりX線吸収スペクトル解析の自動化が可能に ～データの類似度に着目し定量的なスペクトルの解析を実現～

本研究成果のポイント

- 機械学習によってX線吸収スペクトルから物理量を自動で抽出することを実現
- データ解析における解析者の主観を除き、客観的かつ高精度な定量分析に成功
- ノイズの多いデータの解析も可能に。測定効率化や微弱信号解析への応用に期待

【概要】

大学共同利用機関法人 高エネルギー加速器研究機構(KEK)物質構造科学研究所の小野 寛太 准教授と東京理科大学 鈴木 雄太 大学院生(研究当時:修士2年、現所属:総合研究大学院大学高エネルギー加速器科学研究科)は、統計数理研究所の日野 英逸 准教授らと共同で、機械学習を用いて物質・材料研究に必要な不可欠なX線吸収スペクトルの解析を自動化・高効率化する手法を開発しました。

X線吸収分光法(XAS)は、物質・材料の機能と性質を支配する電子状態や化学状態の情報を得ることができる手法で、物質・材料研究において広く利用されています。XAS実験データを解釈し、必要な物理量を取得するためには専門家が目で見て判断する必要がありました。

本研究では機械学習(※1)の一種である多様体学習(※2)および、データの類似度(※3)の概念をX線吸収スペクトル解析に応用することで、XASに内包された物理量を自動的かつ高精度で予測すると共に、大量のXASデータを効率的に解析する手法を開発しました。さらに、スペクトルを比較するための適した尺度を検討することで、ノイズの極めて多いスペクトルからでも物理量を予測できることを示しました。本手法はX線吸収スペクトルのみならず様々な計測に応用することが可能であり、計測データの解析の効率化につながります。さらに、超高速現象や不安定物質の計測など、これまでは解析が困難であった極めて微弱な信号の解析にも適用できると見込まれ、今後の物質・材料研究の加速と新奇な物理現象の理解に貢献します。

この研究成果は、英国の学術誌「npj Computational Materials」に3月29日オンライン掲載されました。

【背景】

私たちの身の回りは電子材料・磁性材料・電池・触媒などの様々な物質・材料に溢れています。物質・材料の機能や性質はこれらの中に存在する電子の振る舞いに支配されており、この電子の振る舞い（電子状態）を詳細に解析することで、より高機能な新物質・材料の開発につながります。

物質中の電子状態を調べるために、物質にX線を照射し、その吸収量を測定するXASが広く用いられています。測定したX線吸収スペクトルを人間が直接理解することは難しいため、データ解析によってスペクトルに対応する物理量（材料パラメータ）を推定し、それを元に議論がなされます。従来は専門家が経験に基づいて、実験スペクトルの微妙な形状変化をシミュレーションや文献と比較することにより、物理量を推定してきました。この方法では解析者の主観が除けない上に、解析作業にかかる人的・時間的コストの高さが問題になります。一方でX線吸収スペクトル計測の効率化は急速に進展しており、現在では24時間で1万本以上のXASスペクトルを取得可能となったため、データ解析の自動化や効率化が強く望まれていました。

【研究内容と成果】

そこで研究チームは、機械学習の基礎技術であるデータの類似度の概念および、機械学習の一種である多様体学習に着目し、X線吸収スペクトル解析への応用に着手しました。材料開発における機械学習の応用はマテリアルズインフォマティクス(MI)と呼ばれる研究領域で、近年急速な発展が続いています。

研究チームは、データ解析における本質的な問題点を議論し、2つのスペクトルがどのくらい似ているのかを示す指標である類似度が定まっていないことから、スペクトル形状の微妙な変化を捉えて物理量を解析するために適した類似度について検討しました(図1)。

この結果、専門家による解析と同程度の精密さで、スペクトルに対応した物理量を予測できることを実証しました(図2)。解析は自動かつ高速に行われ、スペクトル1本あたりの解析時間は0.1ミリ秒程度です。さらに、適切な類似度を用いることで、従来は困難であった非常にノイズの多いデータや分解能の悪い装置で得られたデータの解析も自動で精度良く行うことが可能であると示唆されました。

本研究は、文部科学省元素戦略磁性材料研究拠点およびJST ACT-I(JPMJPR18UE)、CREST(JPMJCR1761)の支援のもと実施されました。

<論文情報>

タイトル : Automated estimation of materials parameter from X-ray absorption and electron energy-loss spectra with similarity measures

著者 : Yuta Suzuki, Hideitsu Hino, Masato Kotsugi, Kanta Ono

雑誌名 : *npj Computational Materials* Vol.5 No.39 (2019)

doi: 10.1038/s41524-019-0176-1

【本研究の意義、今後への期待】

データの類似度の適切な選定は、機械学習をはじめとしたデータ解析の基礎であり、本研究では計測データの類似度の重要性と有用性を示すことができました。本成果によって、X線吸収スペクトルを自動かつ定量的に解析することが可能となり、ハイスループット計測と組み合わせることで、材料開発サイクルの飛躍的な加速につながることを期待されます。加えて、ノイズの多いデータの解析を可能にしたことで、測定時間の大幅な短縮につながるほか、不安定な物質や超高速現象の計測や解析への応用が見込まれ、新奇な物理現象の発見と理解への波及が期待されます。

【参考図】

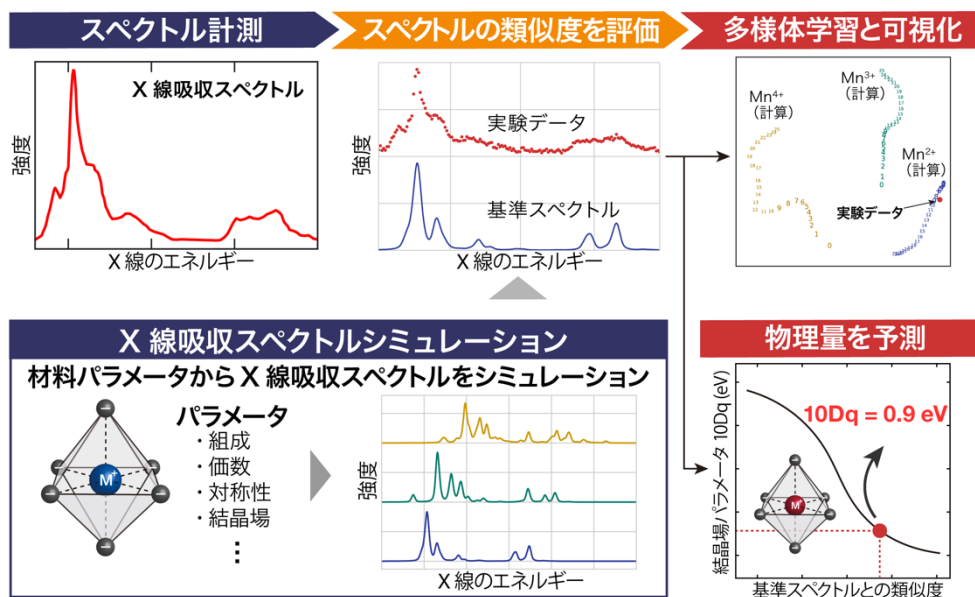


図1. 今回開発した手法の概要

計測したXASスペクトルとシミュレーションとの類似度を評価し、機械学習を用いることにより、化学状態（元素の価数）や物理量の予測をすることが可能となる（それぞれ右端上・下段）

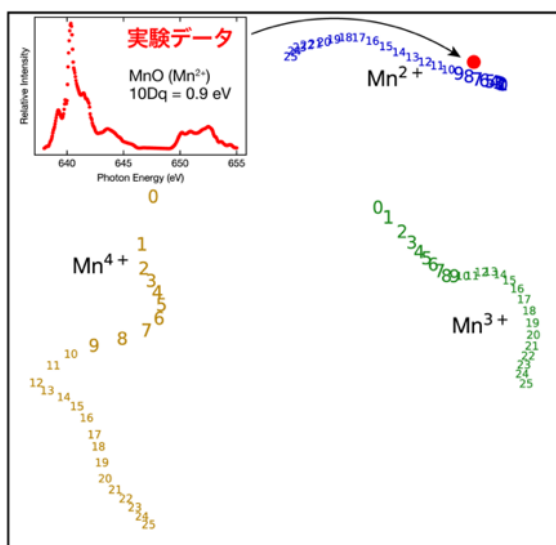


図2. X線吸収スペクトルの類似度の可視化と予測

シミュレーションしたXASスペクトルを多様体学習と呼ばれる機械学習手法で可視化した。図中の各点は1つのスペクトルに対応し、価数（+2～4）と物理量（結晶場パラメータ10Dq（※4）（0～2.5 eV））の変化に対応している。点同士の距離は対応する2つのスペクトルの類似度を示している。

実際の酸化マンガン(II) MnO のXAS計測データを可視化すると、Mn²⁺, 10Dq ~ 0.9 eVに対応しており、専門家の出した結果（Mn²⁺, 10Dq = 0.9 eV）と非常に良い一致を示している。

【お問い合わせ先】

< 研究内容に関する事 >

大学共同利用機関法人 高エネルギー加速器研究機構
物質構造科学研究所 准教授 小野 寛太
Tel: 029-864-5659
E-mail: kanta.ono@kek.jp

< 報道担当 >

大学共同利用機関法人 高エネルギー加速器研究機構
広報室長 引野 肇
Tel: 029-879-6047
Fax: 029-879-6049
E-mail: press@kek.jp

東京理科大学

広報部広報課 (担当: 末永・久保田)
Tel: 03-5228-8107
Fax: 03-3260-5823
E-mail: koho@admin.tus.ac.jp

大学共同利用機関法人 情報・システム研究機構 統計数理研究所
運営企画本部 企画室 URAステーション
Tel: 050-5533-8580
Fax: 042-526-4348
E-mail: ask-ura@ism.ac.jp

【用語解説】

※1. 機械学習

コンピュータ(機械)がデータを学習し、人間が明示的にルールをプログラミングすることなく、データの中にあるパターン・傾向・法則を発見すること。画像認識、マーケティング、自動運転などへの応用が進んでいます。

※2. 多様体学習

元の空間でのデータ同士の距離をなるべく保ったままで、より低次元な空間においてデータを表現する手法。データの次元圧縮ができるため、3次元以上のデータを2次元で可視化するなどの応用があります。

※3. データの類似度(距離関数)

データ同士がどのくらい似ているか、あるいは異なっているかを測る尺度。データが分布する空間における距離を測ることに相当し、このために様々な類似度(距離関数)が提案されています。類似度によって性質が異なるため、目的に応じて使い分ける必要があります。日常生活に例えれば、碁盤の目における2地点間の直線距離と、碁盤の目に沿って進む場合の距離(道のり)を使い分けることに相当します。

※4. 結晶場パラメータ $10Dq$

高温超伝導体や光学材料など物質の多彩な性質を決めているのは遷移金属元素の電子状態です。結晶場パラメータ $10Dq$ は、マンガンや鉄などの遷移金属の電子状態を表すため広く用いられているパラメータです。