

令和2年6月4日

報道関係者各位

大学共同利用機関法人 高エネルギー加速器研究機構
国立研究開発法人 産業技術総合研究所
国立大学法人 総合研究大学院大学
国立研究開発法人 科学技術振興機構

結晶構造解析の自動化

～ブラックボックス最適化により熟練者を上回る解析精度を達成～

本研究成果のポイント

- 数値最適化の応用により X線回折パターン解析を自動化、熟練者を超える解析精度を実証
- 熟練者が1日を要する解析作業を、PC1台で1時間に
- データ解析における解析者の主観を除き、新しい解の候補を発見可能

【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 量子ビーム連携研究センターの小野 寛太 准教授を中心とするKEKおよび総合研究大学院大学の研究グループと、国立研究開発法人 産業技術総合研究所 人工知能研究センターの大西 正輝 社会知能研究チーム長を中心とする研究グループは共同で、数値最適化の一手法であるブラックボックス最適化（※1）手法を用いて、物質・材料研究に必要不可欠な粉末X線回折（PXRD）（※2）パターンの解析を自動化・高効率化する手法を開発しました。

PXRD法は、物質・材料の機能と性質を支配する結晶構造の情報を得ることができる、物質・材料研究において最も広く利用されている分析手法の一つです。PXRDの測定結果から結晶構造の情報を得るためには、リートベルト精密化法（※3）が広く用いられています。この方法には本来の目的である結晶構造情報以外にも多くのパラメータが含まれ、それらの調整に大きな労力が必要とされています。

本研究では、このような状況が機械学習におけるハイパーパラメータ最適化（※4）問題と類似していることに着目し、同問題に対して有効なブラックボックス最適化手法をリートベルト精密化法に応用することで、PXRDパターン解析を効率化する手法を開発しました。本手法を用いることにより、熟練者を超えるフィッティング精度と解

析速度が得られるだけでなく、熟練者がとる典型的な手順では到達できなかった結晶構造の候補を発見することにも成功しました。

本研究のアイデアは、解析結果に影響するパラメータが手作業で調整されている解析手法に応用可能であり、さまざまな分野における計測・シミュレーションデータ解析の効率化が期待できます。さらに、人間の思考の癖や思い込みを排除することで新しい解釈が導かれることも期待され、今後の物質・材料研究の加速と物理現象の理解への貢献が期待されます。

この研究成果は、英国の学術誌「npj Computational Materials」に6月5日オンライン掲載予定です。

【背景】

物質・材料の機能と性質の多くは、結晶構造（原子の並び方）によって決定されます。結晶構造を知ることは物質・材料研究の出発点であり、結晶構造の詳細な解析はさまざまな物理現象の理解や高機能な材料の開発につながります。

PXRD法は代表的な結晶構造解析手法で、物質・材料研究において最も広く利用されている分析手法の一つです。多くの場合、予め仮定した結晶構造や測定装置の特性を盛り込んだ物理モデルから計算されるPXRDパターンが実際に測定されたPXRDパターンに一致するように、モデルのパラメータの更新を繰り返すという手順（リートベルト精密化法）でデータ解析がなされます。この方法によるデータ解析は結晶構造の精密な情報を得ることができる一方で、本来の目的である結晶構造情報以外の多数のパラメータを調整しつつ試行錯誤を繰り返す必要があり、解析作業の人的・時間的コストの高さが問題となります。PXRD測定の効率化が進んだ実験施設では1日に数千件以上の自動測定が可能となっている一方、パラメータの手動調整を伴う解析は熟練者でも1件に1日を要することから、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】

研究グループはかねてより材料データ解析へ機械学習を応用するための研究をしており、PXRDデータ解析の問題が、機械学習においてモデルのチューニングを自動化するハイパーパラメータ最適化問題と類似していることに気付きました。そこで、リートベルト精密化法に対して、ハイパーパラメータ最適化において実績のあるブラックボックス最適化の枠組みを導入することで、自動化する手法を開発しました（図1）。本手法を用いることにより、熟練者を超えるフィッティング精度（測定データとシミュレーション結果の一致精度）が得られるだけでなく、熟練者が1件に1日を要していたデータ解析が、一般的なノートPC1台を用いて1時間程度で行えることを示しました。また、従来用いられてきた熟練者の典型的な手順では到達できなかった結晶構造の候補を発見することにも成功しました。

研究グループでは、物質・材料の研究現場でボトルネックとなっているハイスループットかつ高精度に物質・材料データを取得する課題に取り組んでいます。本研究により物質・材料研究において最も広く利用されている結晶構造解析について、従来は人手に頼っていたデータ解析を完全に自動化することが可能となりました。

本研究はJST未来社会創造事業 探索加速型「共通基盤」領域の研究開発課題「数理科学を活用したマルチスケール・マルチモーダル構造解析システム(研究開発代表者:小野 寛太)」(JPMJMI19G1)で実施されました。また、本研究の成果のうち、ブラックボックス最適化の適用部分は、国立研究開発法人新エネルギー・産業技術総合開発機構(NEDO)の『人工知能技術適用によるスマート社会の実現』事業における「地理空間情報プラットフォーム構築と空間移動のスマート化」で得られた研究成果の一部を適用しました。本研究の一部は、戦略的創造研究推進事業 ACT-I研究領域「情報と未来」の研究課題「物質の結晶構造を高速に予測するデータ解析技術の開発」(研究者:鈴木 雄太)」(JPMJPR18UE)により行いました。

<論文情報>

タイトル: Automated crystal structure analysis based on blackbox optimisation
著者: Yoshihiko Ozaki*1, Yuta Suzuki*2, Takafumi Hawaii*3, Kotaro Saito*3, Masaki Onishi*1, and Kanta Ono*3

*1: 産業技術総合研究所, *2: 総合研究大学院大学, *3: 高エネルギー加速器研究機構

雑誌名: *npj Computational Materials*

<https://doi.org/10.1038/s41524-020-0330-9>

【本研究の意義、今後への期待】

物理モデルによるシミュレーションと測定データの比較は、リートベルト精密化法のみならずさまざまな計測データ解析において基本となる考え方であり、ここで必要とされるシミュレーションパラメータの手動調整のコストは材料研究において共通の問題です。このシミュレーション設定の探索をブラックボックス最適化問題とみなして解くアイデアを提示することは、さまざまな計測データ解析の効率化につながる突破口になると見込まれます。また、単なる解析の効率化のみならず、これまで熟練者の戦略では発見できなかった解の候補を見つけ出せたことから、人間の干渉を減らすことにより解析者の癖や思い込みにより見逃されてきた発見を導くことも期待されます。

本研究成果を活用することにより、結晶構造解析だけでなく、電子顕微鏡、X線顕微鏡、X線吸収微細構造(XAFS)などさまざまな計測機器におけるデータ解析の自動化が可能となり、計測と解析とを統合した計測機器開発に活用されると考えられます。それにより、物質・材料の研究現場で課題になっているハイスループットで高精度な材料データを取得するという問題を解決し、世界最先端の研究開発プラットフォームが構築されることが期待されます。

【参考図】

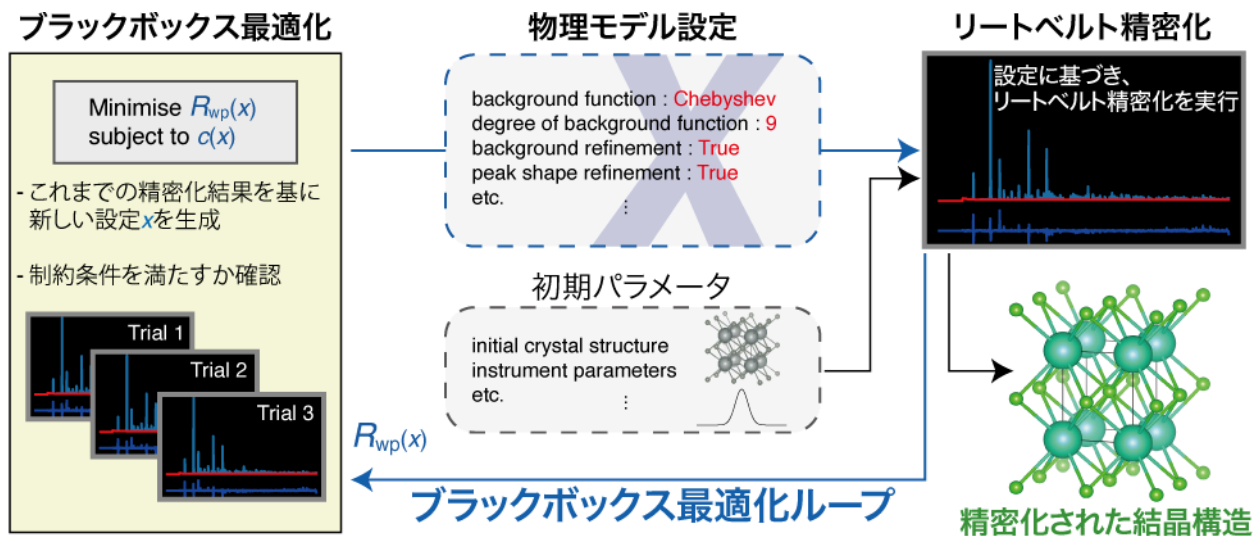


図1. 提案手法の概要

計測した PXR D パターンと、物理モデルから計算される PXR D パターンのあてはまり誤差 $R_{wp}(x)$ を小さくするように、ブラックボックス最適化を用いて物理モデルの設定 x を生成し、あてはまりのよい設定を見つけ出す。物理モデルで設定するパラメータはバックグラウンドを表す関数 (background function) の種類など。

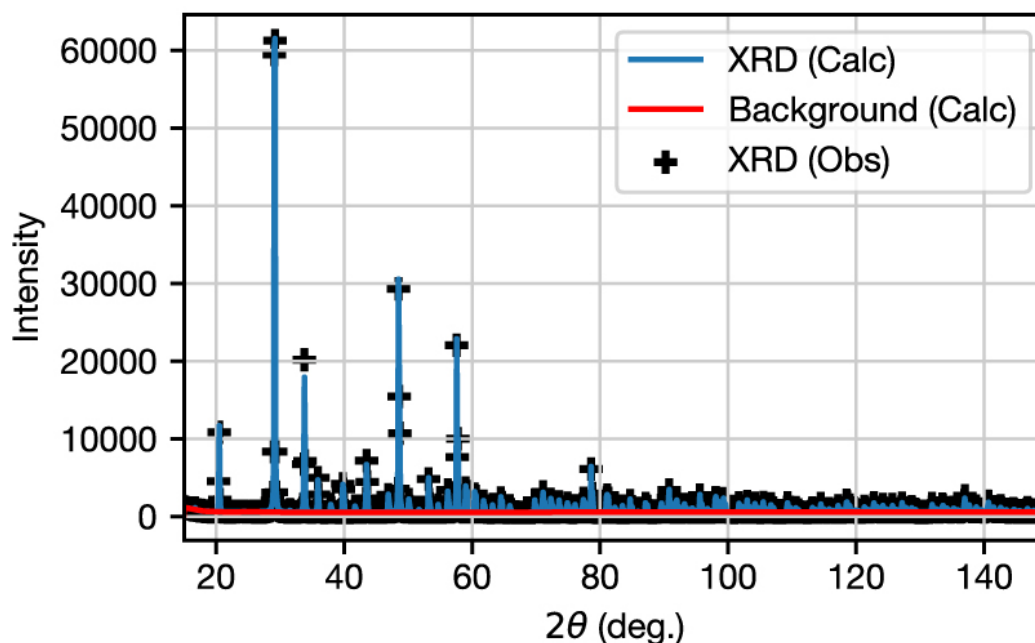


図2. 提案手法を用いた解析結果の例 (酸化イットリウム Y_2O_3 のリートベルト精密化結果)

提案手法により見つけ出した設定による PXR D シミュレーション結果 (青線) は、測定した PXR D データ (十字マーカー) と非常によくあてはまっている。

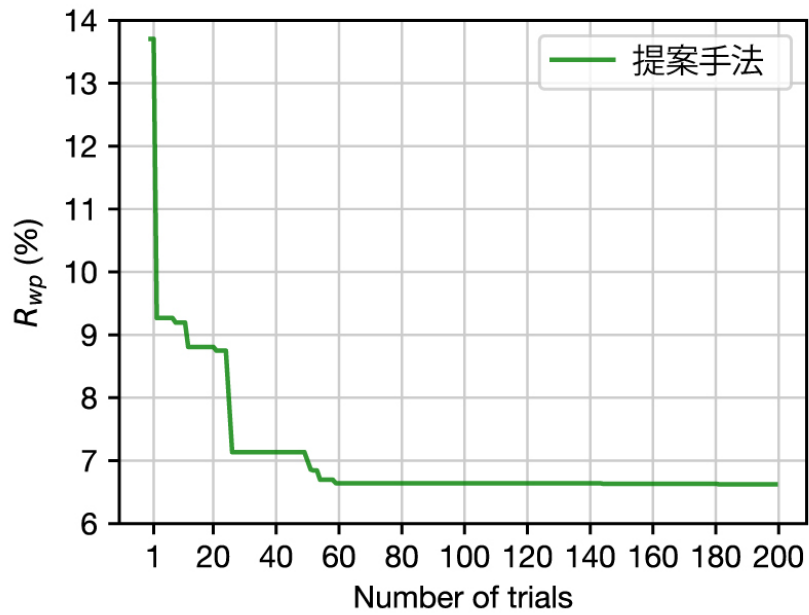


図3. 提案手法による設定探索の履歴の例 (Y_2O_3)

横軸はブラックボックス最適化の繰り返し回数、縦軸はシミュレーションと測定結果のあてはまり誤差 R_{wp} を示す。

あてはまり誤差は小さいほど良く、熟練者では最終的に7%程度になる。提案手法により、平均して100回程度の試行で熟練者より優れたパラメータ設定を見つけ出すことができた。

【用語解説】

※1. ブラックボックス最適化

数理最適化問題のうち、目的関数や制約条件が解析的に与えられないようなものを、ブラックボックス最適化問題と呼びます。ブラックボックス最適化は、機械学習モデルのハイパーパラメータ最適化など、複雑なプログラムやシステムを対象とした実問題において実績があります。ブラックボックス最適化問題を解くためのアルゴリズムをブラックボックス最適化手法と呼び、代表的なものにベイズ最適化などがあります。

※2. 粉末X線回折 (PXRD)

物質にX線を照射すると、X線は物質を構成する原子の周りにある電子により散乱されます。原子や分子が規則的に並んだ物質（結晶）の場合、散乱されたX線は原子や分子の並び方（結晶構造）に応じて回折パターンと呼ばれる独特な強度分布を示します。この現象を利用して結晶構造を調べることができます。X線回折 (XRD) は、一つの結晶（単結晶）を対象とする場合と結晶を砕いた粉末試料を対象とする場合に大別されます。後者は粉末X線回折 (PXRD) と呼ばれ、試料の準備や測定が簡便なことから物質・材料研究において広く用いられる一般的な測定手法となっています。

※3. リートベルト精密化法

多結晶構造を仮定すれば回折パターンを計算することができます。予想される結晶構造と測定装置の特性を盛り込んだ物理モデルからPXRDパターンを計算し、測定から得られたパターンとの差分が小さくなるようにモデルパラメータを繰り返し更新する手順で、PXRDパターンから精密な結晶構造の情報を得ることができます。この手法はリートベルト精密化法と呼ばれ、PXRDパターン解析の主流となっています。

※4. ハイパーパラメータ最適化

機械学習モデルは、一般にユーザーによって設定が必要な幾つかのパラメータを持ちます。このようなパラメータはハイパーパラメータと呼ばれ、機械学習モデルの性能に対する影響力を持ちます。手作業によって、これらのハイパーパラメータを適切に設定するには相応の知識と経験が必要です。ハイパーパラメータ最適化は、最適なハイパーパラメータ設定の探索をブラックボックス最適化問題として定式化し、ブラックボックス最適化手法によって解くことで自動化するものです。ハイパーパラメータ最適化は、作業者の労力削減に加え、性能面においても専門家を上回る設定を見つけ出すなど、近年高い成果を上げています。