

パーシステントホモロジーによる 材料科学データの空間構造解析について

大林一平

東北大学 原子分子材料科学高等研究機構

本講演ではパーシステントホモロジーと呼ばれる位相的データ解析の一手法を用い、機械学習の手法と組み合わせることで材料科学、焼結鈷の X 線 CT の画像データを解析し、その重要な空間構造を捉える試みについて報告する。

位相的データ解析とは数学の「トポロジー (位相)」と呼ばれるアイデアを利用したデータ解析の手法の総称で、データの「形」の情報を定量化することに強みがある。パーシステントホモロジーは位相的データ解析でも特に重要なツールであり、ここ 10 年理論や計算手法に大きな進展があり、アモルファスの解析やセンサーネットワーク、遺伝的進化といった様々な分野で応用が進みつつある。

機械学習は様々なデータから特徴的なパターンを抽出するための手法の総称で、ここ 20 年の理論の発展や入手可能なデータの増大、計算機の性能の向上といった様々な要因で発展を続けている。機械学習はビッグデータと呼ばれるような大規模データでの発展が著しいが、それ以外でも小規模なデータからうまく学習する手法も発展が大きく、また機械学習の手法を手軽に利用するための仕組みも発展してきている。

この 2 つを用いることで、パーシステントホモロジーで定量化した幾何的情報を機械学習の手法で解析することでその特徴的なパターンを取り出すことが可能となる。ここで鍵となるのがパーシステント図と呼ばれるパーシステントホモロジーが持つ情報を可視化した図をベクトル化することで、この変換によってパーシステントホモロジーの情報を機械学習のアルゴリズムに利用することが可能となる。

材料科学では計測によって様々なデータが得られるが、「形」の情報が重要なデータは多い。例えば電子顕微鏡の画像や MD シミュレーションで得られる原子配置情報などである。そこでパーシステントホモロジーと機械学習をそのようなデータに用いることでその形状に関する特徴的なパターンを取り出すことが可能となる。

本講演では焼結鈷の X 線 CT 画像を対象とし、その加熱による組成の分布の形状の変化を上で説明した枠組みによって解析する。焼結鈷は主に酸化鉄と calcium ferrites から構成されており、これらの要素の形状構造をパーシステントホモロジーで解析する。

本講演は高エネルギー加速器研究機構の木村正雄氏と東北大 AIMR の平岡氏との共同研究による成果に関するものである。