

情報計測の高度化に向けた統計的・機械学習的アプローチの今後の展望

東京大学地震研究所・東京大学大学院情報理工学系研究科 長尾大道

統計数理研究所 日野英逸

統計数理の理論的発展と計測技術の発達を両輪として、多様な計測・解析技術と最新の情報科学技術の融合による高度な計測技術が日々うまれており、未知の物質の物理量の測定、新規材料開発とそのハイスループット化の流れはとどまるところを知らない。

折しも人工知能、データサイエンスが盛り上がっており、流行の方法を使ってなにかやりたいけれどもどうしたらよいか分からない、あるいは、データの質や規模に見合わない手法を適用してみただけ、という事例も散見される。科学・産業に寄与する情報計測の高度化の実現には、統計数理の手法の特徴とデータの性質の見極め、大規模データの扱い、結果の解釈性の確保が重要である。また、計測技術の発達により大量データの取得が可能になったものの、その分析は依然人手で行っている、あるいは現状の技術では一回の計測にかかるコストが大きい、といった問題もあるため、単純にデータを集めさえすれば良いというものでもない。

本企画セッションにおいて、伊藤氏は大規模モデルに基づくデータ同化のための新しい4次元変分法を開発した上で、粒成長予測に用いられるフェーズフィールドモデルならびに実際の材料実験で得られた画像データに適用することにより、新規材料開発に資するベイズ推論法を確立した。上木氏は、スパース多項回帰によりマグマの化学組成に基づくテクトニクス場の分類を行った。これまでも機械学習を利用した分類は行われていたが、上木氏の研究は「どの元素が分類に有用であるか、特定のテクトニクス場を特徴づける元素組成はなにか」といった問に答える、地球化学的に大きな貢献である。上野氏は、ガウス過程回帰に基づく能動学習を用い、磁性材料の特性を知るために必要なX線磁気円二色性分光測定の効率化を実現した。磁性材料の特性を定量評価するためにどの程度のエネルギー点数をどのようにサンプリングすれば良いのかは明らかでない。上野氏はスペクトル形状をガウス過程によりモデル化し、事後分散の高いエネルギーにおいて選択的に計測を行うことで測定時間を短縮することに成功した。

新しい計測技術に対応するデータ解析の必要性が統計学の発達の原動力であることは言を俟たず、本企画セッションの講演者のように領域科学の研究者が統計学者と協業しつつ統計的アプローチを身につけて、各々の専門領域における情報計測の高度化を推進していく流れは今後一層加速するものと思われる。