

データ駆動アプローチによる触媒探査

(国立研究開発法人産業技術総合研究所) 永田 賢二

kenji.nagata@aist.go.jp

データ駆動科学とは、大量の実験・計測・計算データから背後に潜む潜在構造や物理法則を抽出することに関して探求する学問であり、データが対象とする学問分野によらず一般的な方法論を議論する学際的な学問分野である。我が国においても、2013年より、科学研究費補助金新学術領域研究「スパースモデリングの深化と高次元データ駆動科学の創成」が発足されてから、生物学・地学分野を中心として、幅広い分野に展開している段階である。

本講演では、こうしたデータ駆動科学の枠組みを紹介するとともに、化学分野にとって関連が深い応用事例をいくつか紹介する。一つは、分光分析により得られるスペクトルデータの解析である¹⁾。スペクトルデータから背後の構造抽出の方策として、ガウス関数などの単峰性のピーク構造を持つ基底関数の重ね合わせで回帰し、得られたピーク位置や強度などから情報抽出する方法があげられる。本講演では、この解析にベイズ推論を適用した事例を紹介する。ベイズ推論とは、ベイズの定理を用いて因果関係を遡ることによりデータ解析を行う枠組みであり、スペクトル分解の場合、ピーク位置などのパラメータのみならずピークの個数までデータから客観的に決定することが可能となる。もう一つは、スパースモデリングによる触媒反応の生成量予測の問題である²⁾。化学における触媒開発の取り組みとして、所望の化学反応に着目し、その反応からできるだけ大量の目的物が生成されるような触媒を探査・開発することである。これは、与える触媒構造を入力とし、得られる生成量を出力として、入出力関係を回帰することで解決できるが、触媒構造からいかにして重要なパラメータ（記述子）を抽出するかが鍵となる。そこで、第一原理計算を利用し、触媒に関する電荷情報や赤外吸収による振動モードに関するパラメータを生成し、これらのパラメータを入力として、スパースモデリングを適用する。スパースモデリングとは、大量の入力に対し、出力との関係と深く関係する入力少数であることを仮定し、データへの適合と少数の関連する入力の選択を同時に行う方法である。これを利用することで、対象の化学反応の生成量に関係する少数の記述子を抽出することが可能となる。

1) Kenji Nagata, Seiji Sugita and Masato Okada, “Bayesian spectral deconvolution with the exchange Monte Carlo method”, Neural Networks, Vol. 28, pp.82-89, 2012,

2) Akira Yada, Kenji Nagata, Yasunobu Ando, Tarojiro Matsumura, Sana Ichinoseki and Kazuhiko Sato, MachineLearning Approach for Prediction of Reaction Yield with Simulated Catalyst Parameters, Chemistry Letters, Vol. 47, No. 3, pp.284-287, 2018.