

研究報告書

「計算論的代数幾何学によるデータ駆動科学の発展」

研究タイプ: 通常型

研究期間: 平成27年10月～平成31年3月

研究者: 永田 賢二

1. 研究のねらい

代数幾何学とは、与えられた多項式が零となる方程式の解により定まる代数多様体について研究する数学の学問である。代数幾何学が現代社会と深く繋がりをもつ学問として、データ解析に関する性能を数理的に解明する統計的学習理論があげられる。統計的学習理論では、与えられたデータをもとに学習のために用意された確率モデルが、データを生成している真の確率分布とどのくらい似ているかを議論するカルバック情報量を導入し、カルバック情報量が最小となる(0 となる)パラメータ空間を議論する必要がある。これは、代数幾何学において、基礎方程式=0 を議論することと対応する。この性質を利用した成果の一つとして、廣中の定理で広く知られる特異点解消により得られる双有理不変量の一つである実対数閾値(RLCT)が、データ解析の予測性能の漸近的挙動を与えることが明らかになった[Watanabe, 2001]。これにより、平成 25 年に新学術領域「疎性モデリング」が発足したことを代表として、多くの分野でデータ解析により科学を推進するデータ駆動科学が創成しつつある。

本研究課題では、こうした背景とこれまでの応募者の研究を踏まえて、ベイズ推定に代表されるデータ駆動科学の方法論構築を行う。また、その方法論構築を通じて、計算機科学・統計学の知見を援用し、高度に発展された計算機を駆使する計算論的代数幾何学とも呼ぶべき学問を推進する。

2. 研究成果

(1) 概要

本研究課題では、代数幾何学を利用して、特異モデルにおけるベイズ推定のアルゴリズム開発を行い、さらに、そのアルゴリズムの結果として得られるサンプリングを利用した代数多様体の系統的構造抽出法の開発を行う。

＜ベイズ推定のアルゴリズム開発、およびデータ駆動科学の応用推進＞

以下 2 つの研究項目を設定し、研究推進を行なった。また、開発したアルゴリズムを利用して、地震動イメージングからの地盤構造推定の問題や、スパースモデリングに基づく触媒探査などの応用に成功した。

【研究項目 1】代数多様体構造の変化に基づく推定の成否判別法の開発

【研究項目 2】代数多様体構造に基づく変数変換による変分ベイズ推定の高精度化

＜MCMC 法を利用した計算論的代数幾何学の推進＞

ベイズ推定を実行する上で有効な技術である MCMC 法によるサンプリングを利用することで、代数幾何学を推進する上で重要な代数多様体の構造抽出や実対数閾値(RLCT)を求

めるアルゴリズムの開発などを行う。現在、論文発表前の時期にあたるため、詳細な記載については省略する。

(2) 詳細

＜ベイズ推定のアルゴリズム開発、およびデータ駆動科学の応用推進＞

【研究項目 1】代数多様体構造の変化に基づく推定の成否判別法の開発

先行研究[Nagata et al., 2012]で開発したベイズ推定によるスペクトル分解の枠組みを拡張し、ピークの個数だけでなくノイズレベルも推定できる方法の開発に成功した[Tokuda et al., 2016]。また、電子や光子のカウンティングにより計測される性質をポアソンノイズとしてモデル化した新たなスペクトル分解モデルを開発した。これにより、計測に要する時間とデータに重畳されるノイズのレベルを明確に関連づけることに成功し、計測時間と推定の成否をベイズ事後分布を通じて判別できる方法を確認した(図 1)。この事後分布の定性的な変化を詳細にみると、多次元のパラメータ空間上での代数多様体構造が明らかに変化していることを見出し、推定の成否とパラメータ空間における代数多様体の両者を関係づけることに成功した。

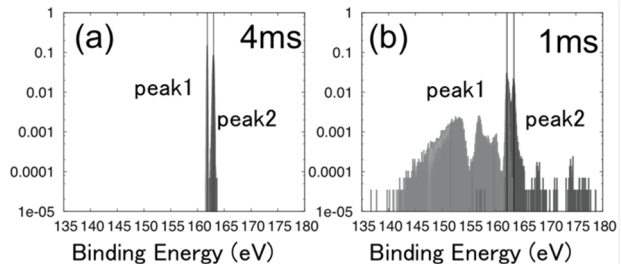


図 1: 計測時間に伴うピーク位置パラメータに関する事後分布の比較

【研究項目 2】代数多様体構造に基づく変数変換による変分ベイズ推定の高精度化

ベイズ推定に変数間などの独立性を仮定した変分ベイズ推定において、独立性による近似が与える推定への影響を明らかにする研究を行なった。具体的には混合正規分布モデルにおいて、MCMC 法によるベイズ推定と変分ベイズ(VB)推定の両者を行い、様々なケースにおいてモデル選択がどのように変化するかを議論した(図 2)。MCMC 法によるベイズ推定は、十分なサンプリングを行う状況下において厳密なベイズ推定の結果と言える。そのため、本研究で得られた MCMC 法と VB 法における結果の違いは、VB 法における近似が与える影響であることが伺える。

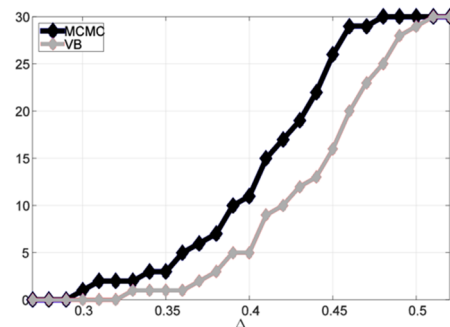


図 2: MCMC 法と VB 法におけるモデル選択結果の比較

3. 今後の展開

データ科学の勃興や人工知能の著しい注目により、データ駆動的アプローチは分野を問わず様々な場面に於て渴望されている。その中でも注目を集めている深層学習や辞書学習などに用い

られる確率モデルは特異モデルであるため、将来的に交換法によるベイズ推定が使われることは十分に考えられる。特に、【研究項目1】をもとに必要な最小限のデータ数についての議論や、【研究項目2】のような代数幾何学に基づいた学習アルゴリズムの評価に代表された数学的に確立される手法を開発し、それをもとに社会貢献できる事例を増やしていくことは学際的な研究が重視される今日において重要な課題であり、こうした取り組みに引き続き研究する予定である。

4. 自己評価

本領域の目標である「数学の力で社会貢献」、および、「その中で数学自身も深化させる」という両面において、十分な成果をあげることができたと考える。データ駆動科学の発展として、地球科学や物性科学、化学分野など多岐にわたる応用を成功させたことは、代数幾何学を基礎とした学習理論に裏付けされた手法の活用であることからまさに「数学の力で社会貢献」という観点に合致する。また、実問題環境において新たな代数幾何学的課題を抽出したり、代数幾何学自身に貢献できるようなアルゴリズムの開発にとりかかれたことは、「数学自身の深化」につながることであると考える。

5. 主な研究成果リスト

(1) 論文(原著論文)発表

1. Kenji Nagata, Rei Muraoka, Yoh-ichi Mototake, Takehiko Sasaki, and Masato Okada, Bayesian Spectral Deconvolution Based on Poisson Distribution: Bayesian Measurement and Virtual Measurement Analytics (VMA), the Journal of the Physical Society of Japan, in press.
2. Yoh-ichi Mototake, Yasuhiko Igarashi, Hikaru Takenaka, Kenji Nagata, and Masato Okada, Spectral Deconvolution through Bayesian LARS-OLS, Journal of the Physical Society of Japan, Vol.87, 114004, 2018.
3. Akira Yada, Kenji Nagata, Yasunobu Ando, Tarojiro Matsumura, Sakina Ichinoseki, and Kazuhiko Sato, Machine Learning Approach for Prediction of Reaction Yield with Simulated Catalyst Parameters, Chemistry Letters, Vol.47, No.3, pp.284-287, 2018.
4. Kano, M., H. Nagao, K. Nagata, S. Ito, S. Sakai, S. Nakagawa, M. Hori, and N. Hirata, Seismic wavefield imaging of long-period ground motion in the Tokyo Metropolitan area, Japan, Journal of Geophysical Research, Solid Earth, Vol. 122, pp.5435-5451. 2017.
5. Satoru Tokuda, Kenji Nagata, and Masato Okada, Simultaneous Estimation of Noise Variance and Number of Peaks in Bayesian Spectral Deconvolution, Journal of the Physical Society of Japan, Vol. 86, 024001, 2016.

(2) 特許出願

研究期間累積件数:0 件

(3) その他の成果(主要な学会発表、受賞、著作物、プレスリリース等)

1. 【招待講演】永田賢二, 「動き出したデータ駆動型計測科学～触媒反応予測とスペクトル解析を例に～」, 第41回日本電子エグゼクティブ交流会, 2018年12月19日.

2. 【企画セッション講演】永田賢二, 「データ駆動的アプローチに基づくキャタリストインフォマティクスへの挑戦」, 企画セッション: 計測インフォマティクス, 第 21 回情報論的学習理論ワークショップ, 2018 年 11 月 7 日.
3. 【学会発表】中山智文, 藤井直樹, 永田賢二, 岡田真人, 混合正規分布モデルにおけるベイズ推定と変分ベイズ推定の比較, 第 21 回情報論的学習理論ワークショップ, 2018 年 11 月 6 日.
4. 【招待講演】永田賢二「データ駆動的アプローチに基づく分析化学への展開」日本分析化学会第 67 年会 産業界シンポジウム-AI, MI 時代への期待と課題-, 2018 年 9 月 14 日.
5. 【プレスリリース】人工知能 (AI) で触媒反応の収率を予測 -キャタリストインフォマティクスで触媒の発見に道一-, 2018 年 1 月 31 日,
https://www.aist.go.jp/aist_j/press_release/pr2018/pr20180131/pr20180131.html