

戦略的創造研究推進事業 CREST
研究領域「計測技術と高度情報処理の融合による
インテリジェント計測・解析手法の開発と応用」
研究課題「ベイズ推論とスパースモデリングによる
計測と情報の融合」

研究終了報告書

研究期間 2017年10月～2023年3月

研究代表者：岡田 真人
(東京大学大学院新領域創成科学研究科
教授)

§ 1 研究実施の概要

(1) 実施概要

戦略目標にある、計測対象の特徴量解析技術の構築の具体例の、シグナル対ノイズ比の低いスペクトルや画像等からの特徴量抽出技術やより少ないデータから有用な情報を引き出す情報再構成技術、異種情報の統合解析技術を構築することを目的とし研究を進めてきた。そのため、ベイズ推論を計測科学に導入したベイズ計測により、計測科学がどう変わるかの具体例を提示し、その具体例の中から、ベイズ計測の情報数理工学的な学理の構築を目指した。ロードマップとしては、中間評価までに、計測限界の定量的評価、系の有効モデルの選択、異種計測の情報統合などが行えることを示し、中間審査後は、情報計測領域への横断的展開をはかっている。以下の四つの課題について研究した。

【課題 1】ベイズ計測:計測限界の定量的評価、有効ハミルトニアンなどの系の有効モデルの選択、異種計測の情報統合などが行えることを示した。計測限界の定量的評価では、スペクトル分解に関して、ベイズ計測による計測限界の定量的評価を行った。有効ハミルトニアンなどの系の有効モデルの選択では、内殻 X 線光電子分光(XPS)のクラスターモデルを生成モデルに組み込んだ、スペクトルデータに基づく有効ハミルトニアンのモデル選択が行えることを示した。異種計測の情報統合に関して、XPS と X 線吸収スペクトル(XAS)の二つのスペクトルを統合することに成功した。特に分光学に集中して、XPS、XAS、メスバウアー分光、XAFS などの研究を行った。そこで得られた数々の成果が認められて、放射光施設 SPring-8 の全ビームラインにベイズ計測を導入するプロジェクトが立ち上がった(岡田 G)。

【課題 2】近似アルゴリズム開発:能動学習によるスペクトル計測の高速化とその最適停止基準の理論的導出による高速化の二重の意味での高速化を実現した。XAS や X 線磁気円二色(XMCD)などのスペクトル計測問題は特定のエネルギー一点で計測した曲線の当てはめ問題と同一視できる。計測点を動的にスパース化することで、従来の 1/5 程度の計測点で同等の計測精度を達成する能動的計測技術を開発した。さらに、その能動的計測をいつ停止すべきかの基準を導出し、Ni 金属 XAS データ解析で実用に供する性能を確認した(日野 G)。

4 次元変分法に対し、2nd-order アジョイント(SOA)法を導入することにより、不確実性評価が可能な手法へと高度化することに成功した。また、シンプレクティック数値解析の考えに基づいて、数値モデルの数値解法に対応する、SOA モデルの超高精度解法を機械的に与える方法論を確立した。また、行列の冪乗法と SOA 法を用いることにより、事後分布の有効自由度の基底を抽出する有効モデル抽出法を確立し、時間依存 Ginzburg-Landau モデルに基づく数値実験を通して、その有効性を確認した(長尾 G)。

【課題 3】モデル構築:畳み込みニューラルネットワークを用いた 3D 単一分子局在型超解像顕微鏡法の開発を行った(日野 G)。

【課題 4】計測データへの適用:ベイズ計測を地球科学に適用することで、有効性を確認するとともに、対象に関する様々な新知見を得ている。地球物理学・地球化学・地質学分野の地球物理学的観測・野外調査・水熱実験・数値計算などの多様なアプローチに対してベイズ計測を組み込むことで、地球コア・マントル・地殻・岩石・火山・地震・海底地形・地球環境などを含む地球全体、特にこれまで数理解析の導入が遅れていた地球物質科学分野に対して先駆的な情報科学と計測科学を融合し、ベイズ計測の浸透に大きく貢献した(桑谷 G)。

2021 年 10 月からは、ベイズ計測を系外惑星の研究に応用する惑星科学ベイズ計測グループ(成田 G)が新たに加わった。本グループは岡田 G とも協力し、独自に開発した多色同時撮像カメラやすばる望遠鏡で得られた時系列データの解析に対してベイズ計測の方法を用いることで、新たな惑星の発見や発見された惑星の特徴量抽出を行った。

また、グループ間連携も有機的に実施できている。岡田 G が開発した計測限界の定量化方法は、長尾 G で実施した異種計測の統合において個別データセットの計測限界の条件下で検証が可能となる。また、日野 G の物理量測定の高効率化においても、計測限界を意識することで最大効率を評価することが可能となる。以上のように、ベイズ計測の情報数理基盤が構築され、さらに桑谷 G、成田 G による実践的なベイズ計測へと展開している。

(2) 顕著な成果

<優れた基礎研究としての成果>

1.

概要:

マルチピークスペクトルをシングルピークの関数の線形和で表現するスペクトル分解に関して、ベイズ推論による計測限界の定量的評価に取り組んだ。MoS₂のSの2p光電子スペクトル(2p-XPS)に対してベイズ推論を計測に適用した。その際に、光電子の放出の統計モデルにポアソン分布を用いた。この枠組みを用いて、2p-XPSの微細なピーク構造まで抽出可能であるかを、定量的に見積もることを実現するためのアルゴリズム開発を行った。

[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)", *Journal of the Physical Society of Japan*, 88, 044003, 2019]

2.

概要:

XAS等のエネルギースペクトル計測の高速化のために、能動学習を用いたガウス過程回帰によるスペクトル当てはめの高速化を検討してきた。これまでは「従来の網羅的計測より少ない計測回数で同等の精度が達成可能」という結果は得られていたものの、逐次的な計測をいつ停止するべきかという問題は未解決であった。そこで、能動学習による当てはめの誤差の理論的な上界を導出し、これ以上計測を続けても誤差程度の利得しか得られないタイミングを客観的に評価する手法を開発した。これにより、外的基準が明確でない計測対象に対しても、逐次的計測による高速化の方法論が適用可能になった。

[Hideaki Ishibashi, Hideitsu Hino, "Stopping criterion for active learning based on deterministic generalization bounds", AISTATS 2020, 386-397, 2020.]

3.

概要:

数値流体力学などの分野で用いられている高次元時系列データ解析手法として動的モード分解(DMD)がある。DMDをベイズ化した手法を開発した。DMDは高次元ベクトルの時間発展を線形作用素の作用で表現し、その固有値、固有ベクトルでデータの時間発展を特徴づける。固有値や固有ベクトルに事前分布を入れてベイズ推定を行い、さらに変分行列分解を用いて効率的な算法を導出した。時系列データ予測や欠損の穴埋めにおいて、ベイズ的時系列予測の標準的な手法と比較して良好な結果を得た。

[Takahiro Kawashima, Hayaru Shouno, Hideitsu Hino, "Bayesian dynamic mode decomposition with variational matrix factorization", *Proceedings of the AAAI Conference on Artificial Intelligence*, 2021]

<科学技術イノベーションに大きく寄与する成果>

1.

概要:

従来のデータ同化では計測データが時系列データであることが仮定され、その適用範囲が計測データとモデルの直接的な比較が可能な場合に限られていた。本研究では、モデル/データ両駆動型データ同化手法を新たに構築することで、非時系列データを含む多様なデータへのデータ同化の適用を可能とする数学的枠組みを創出した。この成果は、これまで限定されていたデータ同化の応用先を飛躍的に広げる基盤技術となることが期待される。

[Tatsu Kuwatani, Hiromichi Nagao, Shin-ichi Ito, Atsushi Okamoto, Kenta Yoshida, and Takamoto Okudaira, "Recovering the past history of natural recording media by Bayesian inversion", *Physical Review E*, 98, 043311, 2018.]

2.

概要:

小角散乱実験で得られたナノスケールの構造情報を示す二次元強度分布データのばらつきを、カーネル密度推定法を適用して抑制する手法を開発した。本手法により短時間の測定で得られた低品質なデータを数倍～10 倍のレベルに高精度化することが可能となり、量子ビーム実験における測定効率を向上させることが可能となった。

[Kotaro Saito, Masao Yano, Hideitsu Hino, Tetsuya Shoji, Akinori Asahara, Hidekazu Morita, Chiharu Mitsumata, Joachim Kohlbrecher, Kanta Ono, “Accelerating small-angle scattering experiments on anisotropic samples using kernel density estimation”, *Scientific Reports*, 9, 1526, 2019]

3.

概要:

双極子間相互作用を含む時間依存 Ginzburg-Landau (TDGL) 方程式は、強磁性体薄膜などのモデルに有用である。本研究では、TDGL に従う 2 次元スピンモデルに対して外部磁場を印加したのち、長時間外部磁場なしで時間発展させた後の 2 次元スピンパターンを新しい方法で分類し、数値時間発展計算を経ないで予測する方法を開発した。

[Anzaki, R., S. Ito, H. Nagao, M. Mizumaki, M. Okada, and I. Akai, Phase prediction method for pattern formation in time-dependent Ginzburg-Landau dynamics for kinetic Ising model without a priori assumptions of domain patterns, *Phys. Rev. B*, Vol. 103, Issue 9, 094408, doi:10.1103/PhysRevB.103.094408, 2021]

<代表的な論文>

1.

概要:

希土類絶縁体化合物の 3d-XPS のクラスターモデルを生成モデルに組み込んだ、スペクトルデータに基づく有効ハミルトニアンとそのパラメータについてのベイズ推論を、交換モンテカルロ法によって実現した。La₂O₃ と CeO₂ の 3d-XPS のクラスターモデルに基づくシミュレーションスペクトルに対して、提案手法を適用したところ、従来物理的考察によって妥当とされていた有効モデルと同じモデルが選ばれることが確認された。

[Yoh-ichi Mototake, Masaichiro Mizumaki, Ichiro Akai, and Masato Okada, “Bayesian Hamiltonian Selection in X-ray Photoelectron Spectroscopy”, *Journal of the Physical Society of Japan*, 88, 034004, 2019]

2.

概要:

本研究では、一般には直接に観測できない合金中の粒構造に対し、データ同化の手法を適用することでダイナミクス情報を抽出することを試みた。合金鋼中の微細な粒構造の静止写真データから粒構造のダイナミクスを形状特徴量の時間発展として抽出し、大規模データ同化と組み合わせることで、成長予測に重要な合金の物性パラメータを推定することに成功した。

[Shin-ichi Ito, Hiromichi Nagao, Takashi Kurokawa, Tadashi Kasuya, and Junya Inoue, “Bayesian inference of grain growth prediction via multi-phase-field models”, *Physical Review Materials*, 3(5), 053040, 2019]

3.

概要:

X 線光電子分光法 (XPS) および X 線吸収分光法 (XAS)には、電子状態に関する補完的な情報が含まれている。つまり、XPS および XAS の統合分析は、ハミルトニアン パラメータを調べるために重要です。この論文では、XPS と XAS を統合するためのフレームワークであるベイズ統合を提案し、推定精度を向上させ、S/N 比の悪いスペクトルから情報を抽出できることを実証した。

[Y. Yokoyama, T. Uozumi, K. Nagata, M. Okada, and M. Mizumaki
“Bayesian Integration for Hamiltonian Parameters of X-ray Photoemission and Absorption Spectroscopy” *Journal of the Physical Society of Japan*,90, 034703, (2021)

§2 研究実施体制

(1) 研究チームの体制について

① ベイズ計測グループ(東京大学大学院新領域創成科学研究科)

研究代表者:岡田 真人(東京大学大学院新領域創成科学研究科 教授)

- ・ ベイズ推論の計測科学への導入
計測科学にベイズ推論を導入する「ベイズ計測」を実現すると共に、それによって、計測限界の定量的評価や、異種計測の情報統合などが行えることを示す。
- ・ スパースモデリングによるベイズ推論の高速近似アルゴリズムの開発
スパースモデリングを用いたベイズ推論の高速近似アルゴリズムを開発する。
- ・ スパースモデリングによる計測対象のモデル構築
スパースモデリングを活用し、複雑な計測対象をモデル化し、そのベイズ推論を実現する。

② データ同化グループ(東京大学地震研究所)

主たる共同研究者:長尾 大道(東京大学地震研究所 准教授)

研究項目

- ・ 4次元変分法データ同化の飛躍的高度化と実問題への応用展開
- ・ 非時系列データ同化の方法論の確立と実問題への応用展開
- ・ モデル/データ両駆動型データ同化法の確立
- ・ 有効状態空間抽出法の方法論の確立

③ スパースモデリンググループ(情報・システム研究機構統計数理研究所)

主たる共同研究者:日野 英逸(情報・システム研究機構統計数理研究所 教授)

研究項目

- ・ 局所性に基づく計測対象のモデル化
スパースモデリングに代表される情報の局所性を利用した統計的データ解析手法を開発する。
- ・ 局所性に基づく高効率な計測の実現
計測における実データ解析の高効率化の実現をする。

④ 「地球科学ベイズ計測」グループ

主たる共同研究者:桑谷 立(海洋研究開発機構海域地震火山部門 グループリーダー代理)

研究項目

- ・ 定常状態画像データからの物理パラメータ・プロセス抽出
- ・ 非時系列画像データからのダイナミクス抽出
- ・ 画像データ、非時系列データからの特徴量抽出を通じた複雑プロセス抽出

⑤ 「惑星科学ベイズ計測」グループ

主たる共同研究者:成田 憲保(東京大学大学院総合文化研究科 教授)

研究項目

- ・ 機械学習による多色トランジット時系列データの最適な高精度解析法の選択
- ・ レプリカ交換モンテカルロ法による視線速度時系列データからの情報抽出
- ・ ベイズ計測による恒星表面の情報抽出

(2) 国内外の研究者や産業界等との連携によるネットワーク形成の状況について

岡田グループでは、情報計測 CREST 3期赤井チームとベイズ的 LARS-OLS(Least Angle Regression - Ordinary Least Squares)法を用いたベイズ推論の高速アルゴリズムに関する共同研究を行なっている。具体的な計測対象としては、超高速分光であるコヒーレントフォノンや EXAFS (extended x-ray absorption fine structure)の研究を行なっており、民間企業1社に紹介・指導を行っている。また希土類化合物の X線光電子分光(XPS)や X線吸収スペクトル(XAS)

の有効ハミルトニアン選択、メスバウアー分光、ベイズ統合、XANES(X-ray Absorption Near Edge Structure)について共同研究を行なっている。これらの実績が認められて、岡田は SPring-8 を運営する公益財団法人高輝度光科学研究センター・放射光利用研究基盤センター 客員研究員に就任し、SPring-8 全ビームラインベイズ化計画を行うことになっている。

長尾グループは、2018年10月から2019年3月にかけて、国内の大手民間会社1社と4次元変分法の応用展開に関する共同研究契約を締結した。また、NASA ゴダード宇宙飛行センターと、データ同化の理論構築とアルゴリズム開発に関する共同研究を計画している。そのほかに、前述の赤井チームとの協働による TDGL 方程式に基づくパターン形成の研究を行っている。この研究では、特にその代表的な例である磁性体中の磁化ダイナミクスについて、実験的に得られた磁区パターンからのパラメータ推定を目指している。

日野グループが顕著な成果<優れた基礎研究としての成果>にて挙げた「能動学習の停止基準」で開発済みの能動学習の停止基準は教師あり学習の枠組みでの理論であったが、例えば逐次的にデータが得られる状況でのカーネル密度推定のような教師なし学習への拡張を行っている。具体的には、小角散乱実験の高速化手法への適用を進めており、赤井 CREST 水牧チームとの共同研究として手法の開発と実データを用いた解析実験を進めている。

桑谷グループにおいては、研究者による共同研究ネットワークとして、東大地震研特定共同研究(B)「機械学習で推し進めるデータ駆動型地球科学の新展開」を実施し、10以上の機関に所属する数十名ほどの若手研究者・学生などとともに機械学習と地球科学を融合する共同研究を推進している。この共同研究では、年3-4回ほど勉強会・ワークショップを開催することで、多数の情報計測融合研究が生まれている。さらに、産業界を含む分野横断型のネットワークを目的として、社会課題であるインフラ維持管理に関わる技術が必要とする産官学の機関が集うインフラ先端技術コンソーシアム(京都大学大学院工学研究科)に、主たる共同研究者の桑谷が幹事として参加している(<https://citi.kuciv.kyoto-u.ac.jp/>)。

成田グループでは、本研究で行っているNASAのトランジット惑星探索衛星TESSと連携した地上観測を通して、国内では主に国立天文台やアストロバイオロジーセンター、海外ではスペインのカナリア天体物理研究所やアメリカのLas Cumbres Observatoryなどとの共同研究を推進している。本研究で運用している多色同時撮像カメラMuSCATシリーズには国内外の研究者から観測依頼が来ており、それを通して多くの国内外の研究者との共同研究に発展している。また、本グループのメンバーは天文学分野においてデータ科学的手法を議論する研究会にも積極的に参加しており、新たなネットワークの開拓にも取り組んでいる。