

小西 貞則 (KONISHI Sadanori)

#### 研究概要

統計科学は、自然科学・社会科学における諸分野の不確実性を有する現象の解明とその本質の探求のため、データから有益な情報やパターンを効率的に抽出・処理するための手法開発と数理的・理論的研究を主な目的としています。

1. 高次元の複雑なデータ構造を探索し、内在する情報を有効に抽出するための手法として、非線形多変量解析という統計手法がある。この非線形多変量解析の研究に関しては、理論や数式に基づく解析的アプローチが難しい複雑な問題が生じる。このような問題に対して、計算機の高度利用を前提とした数値的計算法、特に、ブートストラップ法、EM アルゴリズム、マルコフ連鎖モンテカルロ法と呼ばれる手法によって有効な分析が実行できることを立証し、手法の理論構造を明らかにするとともに、一般理論を構築した。さらに、スプライン、 $B$ -スプラインや動径基底関数などに基づく非線形モデリングや識別・判別、共分散構造分析などに対して、計算機の利用を前提とした新しい視点に立った理論・方法論の研究を行った。

2. 近年の計算機の高度な発展とその利用環境の飛躍的な向上は、データの収集・蓄積・分析を行う上で格段の効率化と高速化を果たしてきた。これにともない、諸科学の不確実性を有する自然現象、社会現象の解明とその本質の探求のための様々なモデルの構築と関連する基礎理論の研究が急速に進展しつつある。データからノイズを除去し、現象の構造を捉えるための有効な非線形モデルの構築は、複雑現象解明にとって極めて重要な役割を担うものである。これらは、スプライン、 $B$ -スプライン、ニューラルネットワーク、カーネル法など様々な方法に基づいて構成されるが、有効なモデルを構築するためにはその構築過程においてモデルの評価が常に本質的となる。これまで統計モデルの良さを評価し選択するための有効な手法として、情報量規準 AIC (Akaike's information criterion) やベイズ型情報量規準 BIC (Bayesian Information Criterion) などが用いられてきた。しかし、複雑な非線形モデリングへの適用に際しては、多くの問題点が指摘されてきた。この問題に対して、情報量に基づくモデル評価基準の研究を行い、AIC や BIC に付与された最尤法の枠を外し多様なモデルの評価を可能とする評価基準を提唱し、非線形モデリングに有効に機能するモデル評価法を提案した。現在、開発した手法を実際問題へ適用するとともに、非線形構造を内在する高次元データに基づくモデリングを研究中である。

3. 確率分布モデルと事前分布を通して得られるベイズ型予測分布モデルは、様々な分野の現象分析のための有効なモデルとして用いられている。このベイズ型予測分布モデルを評価するための情報量規準を構築し、種々の事前分布の探索を可能とする手法を提案した。さらに、汎化能力とモデルの信頼性の向上を図るために、正則化法によって推定したモデルの評価を可能とするベイズ型モデル評価基準の構成について研究した。この評価基準を基底展開法による非線形回帰、識別問題と組み合わせたモデリング手法を提案し、手法の性質と特徴を明らかにした。ベイズモデルは、その複雑さ故に適用上の難点をもつ場合があり、現在、解析的導出を計算機を用いた統計的計算法へ置き換え、可能な限り汎用性の高い柔軟な

手法を開発中である。

4. 電子化された計測・測定技術の進歩は、科学の諸分野で現象過程の連続的な計測・測定と大規模なデータの蓄積を可能とした。この結果、生命科学、システム工学、地球・環境科学などの分野で、複数の特性を経時的に測定した動作過程のデータや時間的に変化する空間データの情報が日々大量に蓄積され、データベースとして組織化されつつある。本研究では、時間の経過にともなって高頻度で観測・測定されたデータや時空間データを関数化処理し、処理した関数化データ集合に基づく非線形現象解明のためのモデリングと理論・方法論の研究に取り組み、識別・判別法、次元圧縮法に関して、いくつかの関数データ解析手法を提唱することができた。開発した解析手法を生命科学、生物工学や気象学などへ応用し、その有効性を立証した。

「マス・フォア・インダストリ」にかかわる H20, 21 年度の研究実績概要

産学連携事業の一環として、新日本製鐵株式会社との共同研究を実施し、鉄をつくる溶鉱炉で精錬された溶鋼を鑄造する製鋼工程において、工程間での溶鋼温度降下量と様々な操業因子を結びつけるモデルを構築して、溶鋼温度を予測・制御する問題に取り組んだ。その結果、蓄積された大規模データを統計的に解析することによって、製造現場で適用可能な溶鋼温度降下量予測モデルを構築できた。この研究成果は、溶鋼温度管理方法、装置及びプログラムとして特許申請することになった。さらに、大量に蓄積されたデータの中から非線形モデリング手法を適用して、製鋼工程における複雑な現象を理解し解明し、予測・制御するための研究に取り組んでいる。

研究業績

1. Ando, T., Konishi, S. and Imoto, S., Nonlinear regression modeling via regularized radial basis function networks, *Journal of Statistical Planning and Inference*, 138, (2008), 3616-3633.
2. Ichikawa, M. and Konishi, S., Constructing second-order accurate confidence intervals for communalities in factor analysis, *British Journal of Mathematical and Statistical Psychology*, 61, (2008), 361-378.
3. Hirose, K., Kawano, S. and Konishi, S., Bayesian factor analysis and information criterion, *Bulletin of Informatics and Cybernetics*, 40, (2008), 75-87.
4. Matsui, H., Araki, Y. and Konishi, S., Multivariate regression modeling for functional data, *Journal of Data Science*, 6, (2008), 313-331.
5. Nonaka, Y. and Konishi, S., Variable selection in logistic discrimination based on local likelihood, *Journal of Japan Statistical Society*, 38, (2009), 431 - 450.
6. Matsui, H., Kawano, S. and Konishi, S., Regularized functional regression modeling for functional response and predictors, *Journal of Math-for-Industry*, 1, (2009), 17 - 25.
7. Kawano, S. and Konishi, S., Nonlinear logistic discrimination via regularized Gaussian basis expansions, *Communications in Statistics - Theory & Methods*, 38, (2009), 1414- 1425.
8. Araki, Y., Konishi, S., Kawano, S. and Matsui, H., Functional regression modeling via regularized Gaussian basis expansions, *Annals of the Institute of Statistical Mathematics*, 61, (2009), 811 - 833.
9. Araki, Y., Konishi, S., Kawano, S. and Matsui, H., Functional logistic dis-

crimination via regularized basis expansions, *Communications in Statistics - Theory & Methods*, (2009), 2944 - 2957.

10. Kayano, M. and Konishi, S., Functional principal component analysis via regularized Gaussian basis expansions and its application to unbalanced data, *Journal of Statistical Planning and Inference*, 139, (2009), 2388 - 2398.

11. Kitagawa, G. and Konishi, S., Bias and variance reduction techniques for bootstrap information criteria, *Annals of the Institute of Statistical Mathematics*, 62, (2009), 209 - 234.

12. Tateishi, S., Matsui, H. and Konishi, S., Nonlinear regression modeling via the lasso-type regularization, *Journal of Statistical Planning and Inference*, 140, (2010), 1125 - 1134.

13. Kayano, M., Dozono, K. and Konishi, S., Functional cluster analysis via orthonormalized Gaussian basis expansions and its application, *Journal of Classification*, 27, (2010), DOI: 10.1007/s00357-010-

14. 川野秀一, 廣瀬慧, 立石正平, 小西貞則, 回帰モデリングと  $L_1$  型正則化法の最近の展開, 日本統計学会誌, (2010), (掲載決定済).

15. Hirose, K., Kawano, S., Konishi, S. and Ichikawa, M., Bayesian information criterion and selection of the number of factors in factor analysis models, *Journal of Data Science*, (2010), (掲載決定済).

#### 講演

1. Kawano, S. and Konishi, S., Semi-supervised logistic discrimination via regularized basis expansions, The 7th World Congress in Probability and Statistics, Singapore, August, 2008.

2. Hirose, K., Kawano, S., Konishi, S. and Ichikawa, M., Selection of the number of factors in Bayesian factor analysis, The 7th World Congress in Probability and Statistics, Singapore, August, 2008.

3. 松井秀俊, 小西貞則, 縮小推定に基づく関数回帰モデリングと変数選択, 統計関連学会連合大会, 慶応大学, 2008年9月.

4. 立石正平, 松井秀俊, 小西貞則, Lassoタイプの正則化法に基づく非線形回帰モデリング, 統計関連学会連合大会, 慶応大学, 2008年9月.

5. 川野秀一, 三角俊裕, 小西貞則, 正則化法に基づく半教師あり識別・判別問題, 統計関連学会連合大会, 慶応大学, 2008年9月.

6. 廣瀬慧, 芹川義和, 小西貞則, 確率的カーネル主成分分析とその応用, 統計関連学会連合大会, 慶応大学, 2008年9月.

7. Kawano, S. and Konishi, S., Semi-supervised learning via regularized logistic discrimination, 2009 Joint Statistical Meeting in Washington DC, August, 2009.

8. Hirose, K., Konishi, S. and Ichikawa, M., A choice of the number of factors and hyper-parameter selection in Bayesian factor analysis model, 2009 Joint Statistical Meeting in Washington DC, August, 2009.

9. Tateishi, S. and Konishi, S., Nonlinear regression modeling via Bayesian regularization with Lasso-type penalties, 2009 Joint Statistical Meeting in Washington DC, August, 2009.

10. 廣瀬慧, 小西貞則, Group Lassoに基づく因子分析モデリング, 統計関連学会連合大会, 同志社大学, 2009年9月.

11. 川野秀一, 小西貞則, 関数データに基づく半教師あり識別・判別問題, 統計関連学会連合大会, 同志社大学, 2009年9月.

12. 立石正平, 小西貞則, 関連ベクターマシンに基づく非線形回帰モデルと変化点探索, 統計関連学会連合大会, 同志社大学, 2009年9月.

#### 学位

理学博士 (広島大学)

#### 受賞歴

- ・2004年9月 日本統計学会賞
- ・2006年8月 American Statistical Association Fellow
- ・2006年9月 応用統計学会優秀論文賞 (共同受賞)
- ・2009年4月 平成21年度科学技術分野の文部科学大臣表彰科学技術賞 (研究部門)
- ・2009年9月 日本統計学会出版賞 (Konishi, S. and Kitagawa, G. (2008). *Information Criteria and Statistical Modeling*, Springer, New York.)

#### その他特記事項

- ・Konishi, S. and Kitagawa, G., *Information Criteria and Statistical Modeling*, Springer New York, 2008.
- ・小西貞則・越智義道・大森裕浩, 計算統計学の方法 - ブートストラップ, EM アルゴリズム, MCMC-, 朝倉書店, 2008.
- ・小西貞則・竹内純一, 統計的モデリング/情報理論と学習理論, 講談社, 2008.
- ・小西貞則, 多変量解析入門 - 線形から非線形へ -, 岩波書店, 2010.
- ・Editor of "Bulletin of Informatics and Cybernetics."
- ・Associate editor of "Annals of the Institute of Statistical Mathematics."
- ・Associate editor of "Journal of Statistical Planning and Inference."
- ・Associate editor of "Journal of the Japanese Society of Computational Statistics."
- ・Reviewer to the Mathematical Reviews.
- ・日本統計学会評議員
- ・応用統計学会評議員
- ・日本学術会議連携会員