

不確実性のモデル化活動グループ

竹村彰通

情報理工学系研究科数理情報学専攻

概要

不確実性のモデル化活動グループでは、確率論、統計学、カオス理論等の手法を用いて、複雑な現実の現象に含まれる不確実性を処理し、ロバストなモデル化の方法論を研究している。現実の複雑な現象は、確定的な理論モデルのみでその挙動を正確に記述・予測することは難しく、何らかの形の不確実性のモデル化が必要である。しかしながら、現象のどの部分を確定的に扱い、どの部分を統計モデルなどにより不確実性として扱うかの切り分けはあきらかではなく、ロバスト性という観点から不確実性のモデル化の総合的な方法論を確立することを目指して、研究成果を蓄積しつつある。

1 研究の目的

現実の様々な現象をモデル化する際に、現象のどの部分を確定的に扱い、どの部分を非確定的に扱うかという切り分けが重要である。複雑なシステムを扱う限りにおいて、様々な不確実性を避けることはできないから、純粋に確定的なモデルではロバスト性に欠ける。一方で、単に不確実性と言っても、不確実性には例えば現象に関する知識の不足等を含め、さまざまな側面があり、確率的なモデルを設定して統計的推論をおこなえばよい、といった単純なものではないことに注意する必要がある。例えば時系列解析一つをとりあげても、統計学で扱う時系列モデルの他にもカオスに代表される非線型システムに基づいた解析を用いることもできる。

このように、モデルにおける確定的な部分と非確定的な部分の切り分け、さらに非確定的な部分

の扱い、のバリエーションを考慮すれば、現象のモデル化において様々なアプローチが可能である。実際多くの競合する方法論について、それらのメリットが個々に主張されそれぞれに研究されているのが現状である。このような中で重要な研究目標は、ロバスト性の観点、すなわち与えられた現象へのモデルの安定的な適合と予測の観点から、多くの方法論を統一的に比較しすぐれたモデルを選びだす指針を与えることである。このような指針を与える基礎研究として、個々の手法の一層の深化が必要であることは言うまでもない。

以下では、このような目標を念頭におきながら、本年度の研究成果の概要と今後の研究の展望について述べる。

2 研究の内容

本グループは「不確実性のモデル化と予測」および「ロバスト時系列計算」の二つのサブグループに分かれており、前者では統計的モデリング及びカオス理論に基づくモデリングを研究している。後者では、揺動散逸原理に基づく時系列解析を研究している。ここでは現在研究が進行中のいくつかの研究テーマについて概要を紹介する。

1. 確率場の理論と応用

近年になって、空間的な拡がりを持ったデータ解析が重要となって来ている。空間データ解析の基礎理論となるのは、時系列解析の添字を多次元に拡張した確率場の理論である。確率場の理論は、確率論と幾何学の境界領域にあり、双方の道具を駆使して新たな展開が得られている。

確率過程や時系列解析の理論は応用上の重要性もあり、すでに大きく発展し確立した分野である。離散時間の時系列解析に加えて、連続時間の確率過程の理論も近年では例えば数理ファイナンスへの応用などの目的から、実際の現象の解析の道具として用いられはじめている。さらに最近になって、地理的なデータの集積とともに空間統計データの解析の重要性が高まって来ている。空間統計データの数学的な基礎は確率場の理論によって与えられるものである。

確率過程から確率場への一般化は、添字集合の1次元から多次元への一般化であるが、添字集合が多次元化するとともに、これまでの確率論的な方法に加えて、幾何学的な手法を併用する必要がある。例えば、さまざまな方向からの断面に関するデータが得られるような場合には、添字集合が球面やグラスマン多様体などの場合も扱わなければならない。このように確率場の理論では、確率論と幾何学の融合が必要であり、解決すべき問題が多く残されている。

ガウス確率場の理論においては、最近になって「オイラー標数法」や「チューブ法」の方法が急速に発展しており、竹村およびその共同研究者によって多くの成果が得られている。特に2003年秋には Taylor, Takemura, Adler によって完全に一般的な設定のもとでオイラー標数法の近似の精度を保証する結果が得られ、永年の懸案であったオイラー標数法の正当化が与えられたことは重要な結果である。

2. ベイズモデルの構築と予測

近年ではベイズ法が実際の有用性から注目を集めている。特に統計的予測においてはベイズ法が有効であり、情報幾何学の手法を用いて、ベイズモデルの予測性能を評価することができる。このなかで、従来の情報幾何学では扱われて来なかったパラメータ空間の大域的な情報幾何学的性質の解明が進んできて

いる。

ベイズ統計学では、Jeffreys 事前分布をはじめとする無情報事前分布について数多くの研究がなされてきたが、現在では、どのような目的にも使える万能な無情報事前分布を、伝統的な主観確率の考え方に基づいて構成することは困難であることが共通の認識となっている。この困難について、駒木は、観測量を与えたもとでの、パラメータと予測すべき量との条件付相互情報量が予測のリスクであることに着目し、条件付相互情報量を評価することによりこの困難を自然に回避できることを示した。これはベイズ予測理論において基礎となる重要な結果であり、今後の理論的な発展が期待できる。

本研究におけるアプローチと密接に関連するものとして、情報理論的な方法により無情報事前分布を構成する Bernardo や Barron 等の研究がある。彼らはコンパクト性に関する人為的な仮定の下で Jeffreys 事前分布を導出しているが、Jeffreys 事前分布が必ずしも良いベイズ推測を与えない多くの例を説明できず、本研究の結果と比較して不十分なものであると考えられる。

3. 離散データ解析における MCMC 法

離散データ解析においてマルコフ連鎖モンテカルロ法 (MCMC) によって正確な推論をおこなう方法が注目されている。理論的にはグレブナ基底を用いた代数アルゴリズムが利用できるが、問題のサイズが大きくなるとグレブナ基底を求めるのは容易ではなく、基底の理論的な考察が重要である。最近になって青木・竹村によりマルコフ基底が極小であるための必要十分条件などが得られた。さらに不変マルコフ基底なども解明された。さらに分割表に関する具体的な検定をおこなうのに必要なマルコフ基底の要素のリストが整備されつつある。これらの結果は計算代数学でも注目されて来ており、統計学と計算代数学の橋渡しとしての役割も期待できる。

4. 複雑システムモデリングの理論と応用

カオス理論や力学系理論に基づく解析法は、伝統的な時系列解析とは異なる新たな視点と方法を提供するものであり、生体や経済現象など複雑な現象に対して有効であることが期待される。合原らはこのような観点から新たな手法の開発に取り組んでいる。より具体的にはマルチフラクタルの観点からの新しい時系列解析の手法の開発、大規模な非線形ネットワークにおける擬アトラクタ間の不規則な遍歴現象の発生要因の解析、実際の神経細胞に似た特性を持つユニットを組み合わせたニューラルネットの数理モデルの開発、などの成果を得ている。

5. 揺動散逸原理に基づく時系列データの非線形構造の解析

揺動散逸原理に基づく時系列解析の分野で、岡部らは揺動散逸定理の連続時間への拡張などの理論的成果とともに、実際の時系列データの定常性・異常性・因果性・決定性を検証する手法を開発している。そしてこれらの手法を地球科学・治水工学・医学・金融工学等への広い応用分野への応用を進めている。

より具体的には、1) KM_2O -ランジュヴァン方程式の揺動過程に対する非線形解析による非線形フィルターの構築のアルゴリズム、2) 時系列の異常性の兆候をリアルタイムで検出する異常性のテスト Test(ABN) の開発、3) KM_2O -ランジュヴァン方程式論に基づく定常性の検証 Test(S)・決定性の検証 Test(D)・因果性の検証 Test(CS) を施すことによる時系列の超ロバストなモデル化、4) これらの手法によって構築された時系列モデルによる将来の予測、という四つの柱について多くの研究成果を得ている。

3 テクニカルレポート

本グループでは、テクニカルレポートの形で研究成果の早期の発信をはかっている。ここでは本

グループのメンバーによる本年度のテクニカルレポートの中から抜粋して、最新の研究成果を概観する。

METR 2003-16

Kei KOBAYASHI, Hideki KAWASAKI and Akimichi TAKEMURA, Parallel matching for ranking all teams in a tournament.

スポーツのトーナメントに参加するチームの順位づけの並列アルゴリズムを提案し、その性質を説明した成果を導出している。

METR 2003-20

Kentaro TANAKA and Akimichi TAKEMURA, Strong consistency of MLE for finite uniform mixtures when the scale parameters are exponentially small.

有限個の一様分布の混合分布において尺度母数を標本数とともに小さくした場合の最尤推定量の強一致性についての結果を証明している。

METR 2003-22

Yo SHEENA and Akimichi TAKEMURA, An asymptotic expansion of Wishart distribution when the population eigenvalues are infinitely dispersed.

ウィシャート行列の母分散行列の固有根が無限に拡散していく状況での、ウィシャート分布の標本固有根と標本固有ベクトルの分布の高次の漸近展開を導出している。

METR 2003-25

Satoshi AOKI and Akimichi TAKEMURA, Invariant minimal Markov basis for sampling contingency tables with fixed marginals.

周辺度数の固定された分割表の集合上のマルコフ連鎖を構成する際に、各要因の水準の入れ換えを変換群としてとらえ、不変なマルコフ基底の諸性質を導出している。

METR 2003-26

Jonathan TAYLOR, Akimichi TAKEMURA and Robert ADLER, Validity of the expected

Euler characteristic heuristic.

Adler によって 1970 年代に提唱されたオイラー標数法の一般的な設定のもとでの正当化は永年の懸案であったが，ここで一般的な有効性の証明が与えられた．

METR 2003-31

Yasunori OKABE and Masaya MATSUURA, On a non-linear filtering problem for multi-dimensional stochastic processes.

揺動散逸原理の観点から，多次元確率過程の非線形フィルタリングの結果を与えている．

METR 2003-32

Yasunori OKABE, On a KM_2O -Langevin equation with continuous time (1).

KM_2O -ランジュヴァン方程式の理論の連続時間への拡張を与えている．

METR 2003-35

Kentaro TANAKA and Akimichi TAKEMURA, Strong consistency of MLE for finite mixtures of location-scale distributions when the scale parameters are exponentially small.

有限個の位置尺度分布の混合分布において，尺度母数を標本数とともに小さくした場合の最尤推定量の強一緻性についての一般的な結果を証明している．

METR 2003-38

Satoshi AOKI and Akimichi TAKEMURA, The list of indispensable moves of the unique minimal Markov basis for $3 \times 4 \times K$ and $4 \times 4 \times 4$ contingency tables with fixed two-dimensional marginals.

2次元の周辺度数を固定した $3 \times 4 \times K$ および $4 \times 4 \times 4$ の3元分割表に対する，マルコフ基底の要素のリストを与えている．

METR 2003-39

Yasunori OKABE and Masaya MATSUURA, Chaos and KM_2O -Langevin equations.

KM_2O -ランジュヴァン方程式とカオス理論の関係についての諸結果を与えている．

METR 2003-41

Yasunori OKABE and Masaya MATSUURA, On an abnormality test for detecting initial phases of earthquakes.

揺動散逸原理にもとづく異常性検出手法を地震データに適用してその有用性を示している．

METR 2003-43

Fumiyasu KOMAKI, Shrinkage priors for Bayesian prediction.

統計的予測問題にベイズ法を適用した場合の縮小型事前分布について幾何学的な観点からの諸結果を与えている．