

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、現実の複雑な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自明ではなく、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。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り、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が重要である。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、空間的な広がりを持ったデータ解析が重要となってきている。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。

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」など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。この結果は確率論に関して最も権威のある *Annals of Probability* に受理され近々出版予定となっている。

2. ゲーム論的確率論に基づく予測理論

本年度はゲーム論的確率論に基づく予測理論の研究において大きな進展があった。ゲーム論的確率論は 2001 年に Glenn Shafer and Vladimir Vovk, “Probability and Finance, It’s Only a Game!”, Wiley, が出版されて以来注目されている。本プロジェクトではこの本の著者である Glenn Shafer 氏 および Vladimir Vovk 氏をそれぞれ 2 週間程度招聘した。これらの招聘以降、竹村は Shafer および Vovk 氏と、ゲーム論的確率論に基づく予測理論について共同研究を開始した。この共同研究からすでに 2 本のテクニカルレポートが生まれており、今後さらなる共同研究の発展が期待できる。ここでゲーム論的確率論について簡単に説明する。

確率とは何かということに関しては、歴史的に多くの哲学的な議論がなされて来たが、1933 年にコルモゴロフによって測度論的確率論の公理的扱いが出版されるとともに、確率は「点」や「線」と同様に、公理的に定義されその意味を考える必要はないものと理解されるようになった。コルモゴロフの測度論的確率論によって確率論の数学的基礎が固められ、数学の一分野としての確率論はその後大きな発展をとげることとなった。しかしながら公理的な扱いは「確率の操作」を「確率の解釈」から分離することに成功しているだけであり、「確率の解釈」を不問にすることによって数学としての確率論の発展を保証したものと考えることができる。

その後も、ベイズ流の主観確率をめぐる議論などで、確率の解釈に関する議論は続いて来たが、コルモゴロフの測度論的確率論によってかわるような確率論の体系は生まれてこな

かった。しかしながら、Shafer and Vovk によるゲーム論的確率論は、賭けとゲームの文脈の中で、事象の確率を内在的かつ整合的に定義できることを示した。さらにこの文脈において、大数法則、中心極限定理、重複対数法則などの確率論の基本的な法則を、測度論を前提とすることなく導出することに成功している。また数理ファイナンスにおける基本公式であるブラック・ショールズ公式の導出にも成功している。このようにゲーム論的確率論はコルモゴロフの測度論的確率論とは全く別に、不確実性を扱うための一つの基本的枠組として非常に期待できる新しい理論体系である。竹村と Sharef, Vovk 両氏の共同研究は順調に進んでいるおり、今後さらに共同研究の範囲をひろげて行く予定である。

3. 統計モデルの情報幾何とベイズ理論

最近、ベイズ的手法の実用的な統計手法としての有効性が、広く統計科学をはじめとする情報分野の研究者に認識されてきている。本研究では、統計モデルに基づいた、性能の良いベイズ予測・ベイズ推測をシステムティックに構成する手法を開発することを目的としている。従来、数理統計学におけるベイズ推測理論ではパラメータ推定に関して多くの研究がなされてきたが、それらの多くは人工的な損失関数にもとづいている。本研究では、ダイバージェンスと情報幾何に基づいて問題を定式化し、パラメータの推定量よりも予測分布や事後分布そのものの評価をおこなうことにより問題の本質的な部分を取り扱っている。統計モデルの多様体が、大域的な幾何学的性質に関するある条件をみたまず場合には、ジェフリーズ事前分布を利用した予測を優越する予測分布が構成できることが示し、さらに個別の重要なモデルに関して、許容性やミニマックス性などの最適性を備えた予測分布の構成法の研究をおこなった。統計モデルに対応する多様体の大域的な性質がベイズ統計理論と直接結びついていることを明確にし、多

様体の大域的な微分幾何学的性質を調べることにより従来良いとされてきた予測（例えばジェフリーズ事前分布に基づくベイズ予測）を優越する予測を構成することを可能にした。従来、情報幾何では局所的な性質を調べれば十分であったが、ベイズ予測分布の構成法を情報幾何学的な視点から考えると、モデル多様体の体積増大度などの大域的性質が本質的な役割を果たす。ベイズ推測理論と情報幾何の特に大域的な性質を本質的に結び付ける視点は世界的にもほとんど見られないものである。

4. MCMC 法による離散データ解析

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。これまで、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、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。また本年度は標本空間の任意の2要素の間の「距離」を常に縮めることができる、という性質をもつマルコフ基底を、ノルム縮小マルコフ基底として定義し、その性質の解明を行なった。これらの結果は計算代数学でも注目されてきており、統計学と計算代数学の橋渡しとしての役割りも期待できる。

5. 複雑システムモデリング

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。合原らは今年度、遺伝子・蛋白質ネットワークや部分放電現象などの実在する複雑システムの数理モデルを構築する

と共に、その複雑現象の背後に存在するゆらぎ特性や放電頻度特性の数理構造を理論的に明らかにした。また、2004年12月6日より8日に、国際シンポジウム “International Symposium on Complexity Modelling and its Applications” を、東京大学本郷キャンパス工学部6号館において、科学技術振興機構 ERATO 合原複雑数理モデルプロジェクトと共催で開催し、内外の研究者を集めて、複雑システムモデリングの理論と応用に関して集中的に議論を行なった。

6. 揺動散逸原理に基づく時系列データの非線型構造の解析

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。

これまで、非平衡統計物理学における揺動散逸定理の数学的構造を調べ、退化した弱定常過程に対する KM_2O -ランジュヴァン方程式論を建設し、マサニ・ウーナ以後未解決であった強定常過程に対する非線形予測問題を解決した。さらに、揺動散逸定理を時系列解析の指導原理として、時系列の定常性・異常性・因果性・決定性を検証する $Test(S)$ ・ $Test(ABN)$ ・ $Test(D)$ ・ $Test(CS)$ を提案し、これらのテストを実行する統合化システムを開発し、このシステムをさまざまな分野の問題に適用している。

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内容を抜粋して紹介することにより、最新の研究成果を概観する。

METR 2004-15

Akimichi TAKEMURA and Satoshi AOKI, Distance Reducing Markov Bases for Sampling from Discrete Sample Space, 18pp., March 2004. To appear in *Bernoulli*.

グレブナー基底を含む各種のマルコフ基底の性質を, 各ステップで状態間の距離を縮めることが可能かという観点から研究している.

METR 2004-19

Hidehiko KAMIYA, Peter ORLIK, Akimichi TAKEMURA, and Hiroaki TERAOKI, Ranking Patterns of the Unfolding Model and Arrangements, 29pp., April 2004.

数理心理学において重要な Unfolding Model に関して現れるランクづけのパターンの集合の性質について, 超平面配置の理論の観点から性質を調べている.

METR 2004-25

Tomonari SEI and Fumiyasu KOMAKI, Bayesian Prediction and Model Selection for Locally Asymptotically Mixed Normal Models, 18pp., May 2004.

モデルのパラメータについて局所漸近混合正規性が成り立つ統計モデルについてベイズ予測法とモデル選択規準を導出している.

METR 2004-36

Kei KOBAYASHI and Fumiyasu KOMAKI, Risk-sensitive Mixed Discrete-Gaussian Networks, 28pp., June 2004.

連続変量と離散変量の双方を含むようなネットワーク決定モデルにおいてリスク感受性を表すパラメータを導入し, リスク感受性の大きさに応じた最適決定を求めるためのアルゴリズムを提案している.

The Game-Theoretic Probability and Finance Project Working Paper No.8. Vladimir Vovk, Akimichi Takemura, and Glenn Shafer. Defensive forecasting, September 2004. Available at <http://www.probabilityandfinance.com> To

appear in *Proceedings of the Tenth International Workshop on Artificial Intelligence and Statistics*, Cowell, R., Ghahramani, Z. eds.

ゲーム論的確率論の枠組を用い, 2 値変数のノンパラメトリックな予測問題の設定で, 予測値と実績値の間の乖離が統計的に検出できないような確率予測方式の構成法を導出している.

METR 2004-47

Hisayuki HARA and Akimichi TAKEMURA, Simultaneous Estimation of the Means in Some Poisson Log Linear Models, 19pp., December 2004.

多変量ポアソン変量の平均ベクトルの推定問題において, ポアソン変量が分割表の形で観測され平均ベクトルが分割表の独立モデルを満たす場合について, 最尤推定量の縮小推定による改良を導出している.

The Game-Theoretic Probability and Finance Project Working Paper No.10. Vladimir Vovk, Ilia Nouretdinov, Akimichi Takemura, and Glenn Shafer. Defensive forecasting for linear protocols. February 2005. Available at <http://www.probabilityandfinance.com>

上に述べたゲーム論的確率論の枠組における確率予測法について, 2 値の場合の結果を, 線形プロトコルにおける多値変数の場合に拡張し, 予測値と実績値の間の乖離が統計的に検出できないような確率予測方式を構成している.