研 究 報 告 書

「統計モデル多様体の普遍的な性質のベイズ予測理論への応用」

研究期間: 平成 20 年 10 月～平成 24 年 3 月
研 究 者: 田中 冬彦

1，研究のねらい

ある未知のパラメータをもつ統計モデルに従って、確率的に次々と発生する一連のデータから、次に出てくる値を推定する場合、従来の方法では一連のデータからパラメータを推定したり予測値と信頼区間を与える事が多い。また、より複雑なモデルの場合には未知パラメータに確率分布（事前分布）を導入して解析するが、事前分布の選択についての理論研究は非独立なモデルでは極めて少ない。

そこで、本研究では、従来の統計的推定を予測という視点から捉え直して、統計モデルの幾何学的な性質に注目して、よりよい予測方法を与える普遍的な理論の構築を目指す。特に、独立同一分布の仮定の下でのベイズ予測に関する最近の理論的な結果を踏まえて、これらを時系列モデルや量子統計モデルに拡張することを中心に据える。

2，研究成果

A. 時系列モデルの研究

私は駒木氏との共同研究で、駒木氏による独立同一分布での優調和事前分布の議論を、スペクトル密度のベイズ推定問題におきかえて、時系列モデルにおける優調和事前分布の理論として展開している。それまでは、p 個のパラメータをもつ AR(p)モデルについて、1つ優調和事前分布が構成できていたが、さきがけ研究をすすめることで、さらに以下の成果が得られた。

1） 直観的に理解しやすいパラメータ表示（偏自己相関係数（PAC）に基く表示）の発見
2） Jeffreys 事前分布がパラメータ領域で積分すると発散するのに対して、優調和事前分布が積分有限、つまり、確率分布として解釈できること
3） 2）の帰結として、優調和事前分布に基いたベイズ推定量の許容性の証明
4） 1）を用いて、2次の AR モデルの今まで見つかっていなかった優調和事前分布の発見

優調和事前分布の存在はリーマン多様体上の正の優調和関数（定数を除く）の存在と同等である。その定義は、一般のラプラス作用素を用いた微分不等式の形で表現されるため、個別の統計モデルに対しては解の存在判定も難しい。また、一意ではなく無数に存在しうる。例えば、AR モデルの場合、根座標系を用いると以下のような微分不等式の解 $h(z)$を求めることに帰着する。

$$\Delta h = \frac{\partial}{\partial z_i}\left(g^{ij} \frac{\partial h}{\partial z_j} \right) + \frac{1}{2g}\frac{\partial g}{\partial z_i}\left(g^{ij} \frac{\partial h}{\partial z_j} \right) \le 0$$

$$g^{mh} = \frac{(1-z_m z_h)\prod_{l\neq h}(1-z_l z_m)\prod_{l\neq m}(1-z_l z_h)}{\prod_{l\neq h}(z_h - z_l)\prod_{l\neq m}(z_m - z_l)}$$

また、上の表示は根座標系を用いておりp次多項式の根として定義されているため、計算上は便利であるが優調和事前分布の直観的な意味がつかみづらい。たとえば3次の AR モデルであればAR座標系では図1のように平面と2次曲面ではさまれた形をしている。高次のARモデルでは高次の曲面ではさまれる形になるため、やはり直観的な意味がつかみづらい。これに比べて、PAC座標系を用いるとパラメータの動く範囲は実数の(−1,1)のp次元超立方体の内部になり、積分の発散収束の判定も平易になる。また実用上も、事前分布を数値計算する上で超立方体の方が扱いやすい。

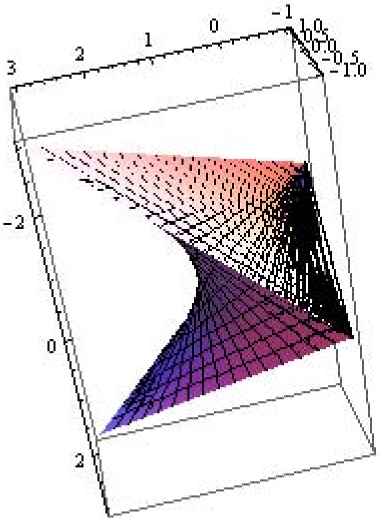

図1：AR(3)モデルの AR 座標系での定常領域（各点が1つの統計モデルに対応）

上述のスペクトル密度のベイズ推定問題を考えることで、独立同一分布の事前分布に関する他の理論的な結果も拡張できる。本研究の成果としては、

5）ARMA モデルのα曲率形式及びα平行事前分布が存在しないことの証明

が挙げられる。α平行事前分布とは、Hartigan (1998)、Takeuchi and Amari (2005) らによって提唱されたものであり、Jeffreys 事前分布が Levi-Civita 接続について平行であるように、α接続で平行な事前分布として定義される。α平行事前分布の存在は統計モデル多様体としての性質に依存して決まる。つまり、リーマン計量だけでなく接続に関する情報も必要である。

本研究では、まず、α平行事前分布の存在条件をα曲率形式の言葉で書きなおした。また、ARMA モデルについてα曲率形式を計算した。α曲率形式の計算では、根座標系を用いた Tanaka and Komaki (2003)にある一連の計算手法がキーになる。その結果、サブモデルであるARモデルやMAモデルはα平行事前分布が存在するが、ARMAモデルでは存在しないことが示された。

B. 量子統計モデルの研究

物理では統計理論というと Fisher 流のパラメータ推定や検定理論が主流で、近年の実験技術の進歩と、小標本での測定精度の向上のために、理論物理、統計理論の両サイドから興味深い

結果が徐々に得られている。しかし、予測分布の理論は理論物理のサイドに浸透しておらず、量子ベイズ予測として扱える問題が理論物理の中で個別に扱われている。本研究では、物理学者が気付きにくいベイズ予測的なアイディアの応用として以下のような新しい成果を得た。

1）波動関数のパラメータ族に対するベイズ予測密度を経由した推定量の構成と積分核を用いた計算公式の導出

2）離散純粋状態族についてのもっとも不利な事前分布の提案

が挙げられる。1）については、有限個の正規直交基底関数を用いて有限次元のヒルベルト空間におとせる場合には、すでに結果が知られているが、ロケーションモデルのような波動関数の中心を平行移動したようなモデルの場合には使えない。多体系の波動関数の場合にはさらに複雑な形になるため、上の手法を用いて実験で得られた測定データから推定する方法が有効である。特に実際の計算においては線型作用素の固有関数を求める必要があるが、解析的には非常に平易な公式も与えた。ここでは、第二種 Fredholm 積分方程式の一般論を用いた。

2）については、統計学特有の問題設定である。量子通信の際に、非直交な純粋状態（例えば波動関数）が複数種類用意されている場合を考える。複数種類のどれが送信されるか不明な状態では、事前分布をどのようにとるべきか。古典のベイズ統計にしたがって、もっとも不利な事前分布(least favorable prior)を提案した。すべての純粋状態が直交する場合には一様分布になるが、非直交の場合には、全く違う状況が起きることを示した。いいかえると、全くわからない場合でも、一様分布を仮定するのは望ましくない。

C. 諸分野との協働について

幾何学的な手法と統計学の接点として、2011 年 7 月に ICIAM にて“Recent development of geometrical approach to statistical analysis”というタイトルにてミニシンポジウムを企画。さきがけの三浦氏（2期生）、伊藤氏（3期生）にも発表していただき、統計理論と「幾何学的な観点」をとりいれて実データを扱っている研究者との交流の場を作った。なお、企画の段階では伊藤氏（北大）も訪問し議論を重ねていた。

またベイズ統計に関しては、駒木氏の科研費と共同で 2009 年 12 月に「ベイズ統計への情報理論的アプローチとその周辺」、2010 年 12 月に「ベイズ統計・量子統計の新展開」を開催し、小規模ながら参加者間で活発な議論をすることができた。

量子統計に関しては、統計学会の周辺での開催がほとんどなく、量子情報・量子計算などの研究集会の中で個人単位で発表を行っているのが現状だった。そのような状況下で、2010 年 4 月に主に理論物理・実験物理で統計に興味がある人が交流・情報交換するためのメーリングリストとして q-stats が発足した。私は物理分野との協働を模索しており、q-stats のコアメンバーとして物理や数学出身の若手と積極的に交流してきた。広く量子情報に携わる学生・若手研究者の交流のための研究集会である、関東 Student Chapter にて、2回ポスター発表を行い、逆に自分でもミーティングの機会を作ってきた。 これらは個人の学術的な業績にはならないものの、結果として、統計関連学会連合大会（通称：連合大会）にて初の量子統計セッションを開催するという一歩を踏み出すことができた。

4, 今後の展開

時系列モデルについて微分幾何学に基いた研究は、数学的にも応用的にも課題が多い。ARモデルであれば理論的には優調和事前分布に基いたスペクトル密度の推定量をもとに、1期先予測を構成することができる。これらの予測方法の数値的な検証が必要になってくる。また、計量経済学における多変量時系列モデルを用いた解析や、最近、空間統計学で利用されている空間ARモデルなどの理論的な解析にこれまでの結果を拡張することも考えられる。

量子統計モデルについては、数学的技巧を別として、古典統計理論の枠に入る概念の数学的な拡張(可換な量を非可換にうつす)は、ある程度、進められてきた。しかし、その一方で、波動関数のパラメータ族のような統計モデルの場合には、古典的な対応物が存在しない。より概念的な考察と整備が必要になり、それこそが統計学の根本である。事前分布の選択やはずれ値に対するロバストネス、推定量の許容性などは、統計理論の研究者にはなじみがあるが、物理学者にはなじみのない概念である。量子論の枠組みの中に、こういった概念をうまく組み込んで定式化するためには、実験的な状況をある程度理解している研究者との議論が必要である。残念ながら、さきがけ終了後は公のサポートもないものの、q-stats の中で交流・議論を通じて、豊かな成果を出していく予定である。

交流の場の提供という意味では、震災で1年ずれてしまったが、2012 年 7 月開催の 2nd ims-APRM(統計の国際会議)で海外の研究者も交えた量子統計セッションを企画している。また、2012 年度の RIMS 研究集会提案「量子論における統計的推測の理論と応用」が無事に採択。q-stats のメンバーに統計理論のチュートリアルと研究発表を 10 月下旬ごろ開催予定である。

4, 自己評価

すべてが手探りであり、先行研究をマイナーチェンジするというよりは、新しいパラダイム・手法を提案する方向での研究だった。今から思うとかなりリスキーであったかもしれない。常套手段があるわけでもなく、手探りのため、細かい点については、いろいろな方針をたてて試行錯誤しており、途中でとん挫したり、行き詰まった部分も多い。それでも、たとえば、AR モデル多様体については、ある程度、まとまった結果が得られたので良かったように思う。量子統計に関しては、数学的な拡張にしても、新しいアイディアの提案についても、統計・物理の両方の分野から否定的な評価を受けてなかなか論文として publish できずにいる。こちらは、q-stats の中で発表・議論を通して論文を改善していく予定である。現時点で見れば、苦労した割に形になった業績が少ないという感も否めない。個人の研究業績に関しては、最終年度に国内外の研究集会で発表したものを今後、論文の形で publish することが重要である。

一方で、分野間連携という観点では、当初の予想以上に大きく進展した。これはさきがけ領域会議などを通じて、他分野と積極的に連携をすすめている研究者に影響を受けたところが大きい。初めは、実験物理との直接連携がしっくりこない理由がわからずに苦労したが、実験物理と理論物理の間にも壁があることを知り、「まずは統計に興味ある理論物理の人と連携すればいい」という方針を得てから、連携のビジョンが見えるようになった。また、初めから統計の A さんと物理の B さんをすぐに協働させるようなお見合い方式でなく、「いつでも誰でも協働できる土壌づくり」を念頭において定期的に会うようにしたのが、大変うまくいっている。(「協働(=結婚)」のプレッシャーがない! 気が合う人をじっくり探せばいい。) 私の側からすれば、理論物理の人から、実験の人が実際にやっていることや関連文献を教えてもらうだけでなく、最近では統計学と

量子論が混ざった深い議論もできるようになってきた。その他、q-stats 以外にも、セッションや研究集会の企画など、これまで全く経験のないことも積極的に行った。海外の研究者の招へいなども、慣れない作業で時間もとられたが、非常に良い経験を積むことができた。今後は、こういった場を継続しつつ、具体的な研究成果を出していくことが重要である。

5．研究総括の見解

時系列モデル、量子統計モデルにおいて従来の統計的推定を予測という観点から捉え直すと共に、より幾何学的性質にも着目して、優れた予測方法を与える手法をベイズ予測の最新の結果を用いて大きく進展させたことは評価できる。

とくに物理学者と継続して協働できる場 q-stats を作り、量子統計モデルの今後の基盤作りにも成功したことも高く評価できる。

本事業に参加することにより、その意識を変え、さきがけスピリットを最もよく体現した研究者の一人であると思う。垣根を越え、横断的な姿勢で研究することはたやすいことではなく、それを実施した田中氏の今後に大いに期待する。

6．主な研究成果リスト

（1）論文（原著論文）発表

1. TANAKA Fuyuhiko and KOMAKI Fumiyasu, "Asymptotic expansion of the risk difference of the Bayesian spectral density in the autoregressive moving average model", Sankhya Series A, Indian Statistical Institute, Vol.73-A (2011), pp. 162－184.

2. TANAKA Fuyuhiko, "Bayesian estimation of the unknown wave function" submitted to Physics Letters A.

（2）特許出願

特になし

（3）その他の成果（主要な学会発表、受賞、著作物等）

1. TANAKA Fuyuhiko, "Application of differential geometry to Bayesian estimation of spectral density", ICIAM 2011, Vancouver, BC, CANADA, July, 2011.
2. TANAKA Fuyuhiko and TAKEUCHI Takuma, "Hypothesis testing of a maximally entangled state under the unknown unitary process", The 14th Workshop on Quantum Information Processing, Montreal, CANADA, December, 2011.