

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成 24 年 6 月 8 日現在

機関番号：12601

研究種目：若手研究(B)

研究期間：2008～2011

課題番号：20700250

研究課題名（和文） 古典・量子相関をもつ系での微分幾何に基づいたベイズ予測理論の研究

研究課題名（英文） Differential geometrical approach to Bayesian prediction theory to classical and quantum correlated systems

研究代表者

田中 冬彦 (TANAKA FUYUHIKO)

東京大学・大学院情報理工学系研究科・助教

研究者番号：90456161

研究成果の概要（和文）：

時系列解析で良く使われる AR モデルで優調和事前分布の便利なパラメータ表示を与えて理論的な結果だけでなく数値計算での有効性も示した。また、ARMA モデルでは微分幾何を用いてアルファ平行事前分布がジェフリーズ事前分布以外に存在しないことを証明した。波動関数の統計モデルで無情報事前分布を定義しミニマックス定理を証明した。波動関数の最適な推定方法を示し、具体例で最尤推定に基いた従来の方法よりも一様に推定精度があがることを示した。

研究成果の概要（英文）：

We obtain a useful parametrization of superharmonic priors in AR models, which are often used in time series analysis. It is shown to be useful not only in theoretical development but also in numerical analysis. We show that there is no alpha parallel prior except for the Jeffreys prior in the ARMA model by using differential geometry. We give a general definition of noninformative prior in a statistical model of wave functions and prove a minimax theorem with respect to this prior. We also derive the optimal estimation of wave functions and give an explicit example where our estimator has a uniformly better performance than that based on MLE.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2008年度	600,000	180,000	780,000
2009年度	700,000	210,000	910,000
2010年度	500,000	150,000	650,000
2011年度	500,000	150,000	650,000
年度			
総計	2,300,000	690,000	2,990,000

研究分野：統計理論

科研費の分科・細目：情報学・統計科学

キーワード：

統計的予測、情報幾何、時系列解析、

量子情報、事前分布

1. 研究開始当初の背景

本研究では、従来の統計的推測を予測という視点からとらえなおして、時系列解析から量子情報科学まで適用できる普遍的な理論の構築を目指す。

まず、本研究で注目するベイズ予測、ベイズ予測分布について簡単に説明しておく。ある未知のパラメータをもつ統計モデルにしたがって次々に発生する一連のデータを考えよう。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て予測値と信頼区間を与えるにすぎなかった。本研究では予測値の分布（予測分布と呼ぶ）を推定する問題を考える。このような考え方は非常に自然である。

たとえば、イベントでの4日間の来客人数が、22人、34人、18人、30人と続いた時に、翌日に配置する店員の数を考えたいとする。単なる予測値(26人)よりも、予測分布（来客人数の分布）がある方が判断しやすいであろう。特にサンプルデータ数が少ない場合には、パラメータの初期分布（事前分布; prior）を仮定して、データに基いて更新した分布（事後分布）で平均化することで予測分布を構成する。これはベイズ統計に基いておりベイズ予測分布と呼ばれ工学諸分野で広く使われている。

パラメータの事前分布が既知の場合にはベイズ予測分布が最良であることが古くから知られており、Aitchison (1976) や Geisser (1993) によるテキストもある。そのため、画像処理や学習理論、ベイジアンネットワークといった様々な分野でベイズ予測分布の方法が暗に陽に应用されている。

しかし、パラメータの事前分布が未知な場合に、どのような事前分布を仮定すればよいかについて、理論的な研究はまだまだ少ない。ベイズ予測分布では事前分布の選び方で大きく結果が異なることもあるため、望ましい選び方の指針が必要である。事前分布の選び方はベイズ統計で古くから知られている難しい問題であり、理論的な提案は Jeffreys 事前分布や Bernardo(1979) の事前分布 (reference prior) など数えるほどである。

研究開始当初は、日本の駒本文保氏やアメリカの George, Lian, ら Yale 大学のグループによる、ミニマックス予測や許容性などの理論的な結果がある程度でているといった状況だった。

2. 研究の目的

このように、研究開始当初から、予測分布の理論的な研究はさほど多くなかったし、現

在も同様である。独立同一分布 (i.i.d.) の仮定の下では、それでも、特定の性質のよいモデルや、大標本では、優調和事前分布を用いることで Jeffreys 事前分布に基いた予測分布を一様に改善できることが駒木氏によって示されている。しかしながら、時系列データのように、一般に独立同一分布の仮定が成立しないような場合は先行研究がほとんど皆無であった。

また、予測分布の理論を数学的に眺めてみると、既存研究での方法論や手法は「次に出てくる値の予測」という考え方を越えた応用にも適用できる。そこで、私は i.i.d. の下で調べられてきた予測分布の構成方法に関する理論を (1) 時系列解析、(2) 量子情報科学に拡張・応用することを目指す。具体的には以下のようになる。

(1) 時系列解析

時系列データ解析における自己回帰モデル (AR モデル) の場合には、優調和事前分布を利用することで精度よく予測できることが理論的には示されている。しかし、筆者達による既存の提案は複素数を経由しており数学的には美しい形だが、ややわかりにくい。そこで、時系列データのベイズ解析を行う実務家に向けて、より使いやすい形・表示の優調和事前分布を提案する。

さらに、応用上の観点から有限の長さの時系列データで実際に理論がどの程度有効か詳細に調べる。

また、AR モデルより広い統計モデルである ARMA モデルの場合について、理論的な正当性をもつ無情報事前分布が存在するかどうかを明らかにする。このような性質は統計モデル多様体の幾何学的な性質で決まるため、そのような事前分布の存在条件を微分幾何学の言葉で与え、具体的な構成方法の提案を試みる。

(2) 量子情報科学

量子系の場合にはベイズ予測の手法自体がまだ整備されていない。物理、特に量子情報の文脈では、個別の量子統計モデルに関連する結果が既にあるので、それらを踏まえて一般論を構築する。

また、古典系でのベイズ予測と事前分布の関係について、理論的な結果を検討し、量子統計モデルにおける望ましい事前分布を提案する。そのような事前分布が必ずしも存在するとは限らないため、存在条件をできるだけ一般的な形で明らかにする。

3. 研究の方法

(1) 時系列解析

まず、AR モデルについては、既に得られている優調和事前分布について、実際に数値的な実験とスペクトル密度の推定を行ってみる。

また、上に述べたように、ベイズ予測分布では事前分布の選び方で大きく結果が異なることもある。理論的な提案は Jeffreys 事前分布など数えるほどである。ここでは統計モデルに対する幾何学的な見方が有効であり、私は駒木氏による微分幾何的な手法をさらに推し進めて、ARMA モデルやその部分モデルでの、無情報事前分布の定義と理論的な正当性について検証する。

(2) 量子情報科学

量子系でのベイズ予測の枠組みは、私と駒木氏が 2005 年に提唱したばかりである。そのため、応用以前に、まだ理論的なフレームワークに関する部分も整備されていない。上にも述べたように、物理の文脈では個別のモデルに関する理論的な結果が多いため、それらの結果を詳細に検討し、一般の量子統計モデルに通用する理論を体系化していく。

ベイズ統計における Jeffreys 事前分布の特徴づけは複数あるが、ベイズ予測に関連する所では、Clark and Barron (1994) による導出が重要である。ミニマックス符号化、情報源の推測に関するゲーム理論的な枠組みで事前分布を考察することができるため、これらの理論を量子論特有の問題に結び付けて定式化する。

4. 研究成果

研究開始当初は、時系列モデルで理論的な正当性をもつ無情報事前分布は 1 次の AR モデル以外にない状況だった。現在、筆者達の提案する優調和事前分布以外にも、若干の提案はあるが、それでも、本研究の成果 (1) ~ (4) は国内外で他の追随を許さない先駆的な研究である。応用上は (1)、(2) の成果が重要で、これはさらに推し進める必要がある。統計理論的には (3)、情報幾何学 (微分幾何学) との関連からは (4) が興味深い成果である。

また、ベイズ予測の量子情報への応用としては、(5)、(6) の成果が得られた。量子ベイズ予測は筆者と駒木氏が 2005 年に提唱した、まだ歴史の浅い理論である。物理の文脈では、具体的な統計モデルを立てて誤差の下限を計算する研究は多いが、本研究は

一般的な量子統計モデルで密度作用素や波動関数のベイズ推定を取り扱っている。特に (6) に関しては、物理 (量子論) とベイズ統計学双方から見て非常に重要な成果である。

(1) AR 過程での有限長データでの優調和事前分布に基いたベイズスペクトル密度の推定方法の有効性の検証

数値的に AR 過程から有限長データを発生させ、優調和事前分布に基いたスペクトル密度の推定の性能を調べた。実データへ適用する場合には、事後分布の構成において、AR パラメータをそのまま用いた単純な棄却法は効率的ではない。点推定と違い、スペクトル密度の推定量はパラメータの取り方に依存しないため、本研究では、パラメータとして偏自己相関係数を利用することを提案した。実際、効率よく計算できることが示された。幾何学的には、これは、座標変換で別の座標系に切り替えることに相当する。

有限長データでの問題点として、定常性の境界付近のパラメータが選ばれた場合、数値的な精度が悪くなることが明らかになった。非定常に近い時系列モデルでは、数値的に発生させた時系列データの長さが短い場合、その挙動が不安定になることが主な原因である。

(2) 偏自己相関係数パラメータを座標系とみなした AR モデルの解析

偏自己相関係数による座標系 (PAC 座標系) を用いると 2 次の AR モデルは Fisher 計量が対角化される。このことを利用して、既存のものより広いクラスの優調和事前分布が見つかった。特に偏自己相関係数をパラメータとする一様分布も優調和事前分布になる。本研究により、理論的な正当性をもつ時系列モデルの無情報事前分布が AR モデルで提案されたことになる。ベイズ法による時系列解析のひとつの指針となることが期待される。

(3) 許容的なスペクトル密度の推定量

Jeffreys 事前分布が improper であるのに対して、優調和事前分布は PAC 系で記述することで、proper になることも容易に示される。したがって、理論的な帰結として、優調和事前分布に基いたスペクトル密度の推定量は Kullback-Leibler 損失の下で許容的であることも示された。

(4) ARMA モデルにおける α 平行事前分布の存在判定とドラムコホモロジーでの定式化

当初の目的としては、ARMA モデルの場合について、優調和事前分布が存在するかどうか、微分幾何学的な量に注目して明らかにしていくということだった。しかし、そもそも ARMA モデルでの事前分布に関しては、理論的な正当性に基く議論は皆無だったため、優調和事前分布にこだわらず、一般的に事前分布の有無を検討する方向にシフトした。本研究では、Hartigan(1964, 1965)と Takeuchi and Amari (2005) で独立に提唱された α 平行事前分布を定常ガウス過程に従う時系列モデルについて導入し、既存結果をうまく適用することによって、存在条件を微分形式に基づいて与えることができた。その結果、AR モデル、MA モデルでは α 平行事前分布が存在する。一方 ARMA(p, q)モデル($p>0$, $q>0$)では α 平行事前分布は $\alpha = 0$ (Jeffreys 事前分布に相当) の場合を除いて一般には存在しえないことが示された。

(5) 波動関数のパラメータ族に対するベイズ推定

本研究では波動関数のパラメータ族が与えられた下での推定を考える。これらは純粋状態モデルとも呼ばれる。損失関数を Bures 距離にとり、測定を固定した下での最適な推定量 (ベイズ波動関数と呼ぶ) を求める公式を陽に与えた。これは量子論特有の問題である。有限次元ヒルベルト空間での問題に帰着できるケースは古くから多くの研究があったが、たとえば、ロケーションモデルのような場合は帰着できない。このような無限次元ヒルベルト空間の場合に、一般的な取り扱いには本研究が初めてである。物理分野では、有限データでの波動関数の推定を未知パラメータの推定に帰着する方法しか考えてこなかった。しかし、具体的な例で最尤推定量に基いた波動関数の推定が非許容的であることが示される。具体的にはベイズ波動関数によってリスクが一様に改善される。

(6) 純粋状態モデルでの無情報事前分布の提案とミニマックス定理

本研究では、古典ベイズ統計での Bernardo(1979)のアイデアをより深く掘り下げていくことで、新しいタイプのゲームとそれに付随する事前分布 (MDP prior と呼ぶ) を提案できた。無限次元ヒルベルト空間であってもパラメータ領域がコンパクトであれば MDP prior は存在する。コンパクト性を外すと存在しない例が有限次元ヒルベルト空間のケースで構成できる。

本成果により、事前分布の観点から「情報がない」ということの量子論的な意味が明確

になった。具体的には以下のような問に対する一つの解答を与えたことになる。(解答は MDP prior, 詳しくは F. Tanaka, 2012, PRA を参照)

問: あるミクロの系 S が、3種類の非直交な波動関数 a, b, c のうちのどれか一つで記述されるとする。我々は、ミクロの系 S がどの状態にあるか全くわからない (情報が無い)。このとき、 a, b, c に対して一様分布 (つまり、それぞれ $1/3$) をふるべきか? もし、一様分布以外ならどのように定めるべきか?

古典論に帰着される場合、つまり、 a, b, c が互いに直交する波動関数ならば、対称性から、 $1/3$ 、つまり等分配が自然な解になる。しかし、量子論が入るとそのような定義は問題が生じる。本研究では、おそらく世界で初めて、このような非常に素朴かつ量子論特有の問題を取り上げ、ベイズ予測の考え方に基いて一つの解答を与えた。

統計理論的にも (5) のような純粋状態モデルでベイズ推定を行う上で、無情報事前分布を選ぶ指針が初めて与えられたことになる。具体的なモデルでのその他の事前分布との比較など豊富なテーマが上の結果から出てくる。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 5 件)

(査読あり)

[1] TANAKA Fuyuhiko, "Noninformative prior in the quantum statistical model of pure states," Physical Review A, American Institute of Physics, Vol.85 (2012), 062305. DOI:10.1103/PhysRevA.85.062305

[2] TANAKA Fuyuhiko, "Curvature form on statistical model manifolds and its application to Bayesian analysis", Journal of Statistics Applications and Probability, Natural Sciences Publishing, Vol.1 (2012), pp. 35-43.

<http://naturalspublishing.com/ContIss.asp?IssID=33>

[3] 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.
DOI: 10.1007/s13171-011-0005-1

(査読なし, 紀要)

[1] TANAKA Fuyuhiko,
"Superharmonic priors for autoregressive
models", Mathematical Engineering
Technical Reports, University of Tokyo,
2009-18, pp.1–20, 2009.
<http://www.keisu.t.u-tokyo.ac.jp/research/techrep/2009.html>

[学会発表] (計 1 1 件)

[1] TANAKA Fuyuhiko,
"Curvature form on statistical model
manifolds and its application to Bayesian
analysis", ICM 2012, Al Ain, UAE, March,
2012.

[2] TANAKA Fuyuhiko, "Noninformative
prior in quantum pure states model."
NEXT2012Nara, Nara, JAPAN, March, 2012.

[3] TANAKA Fuyuhiko,
"Application of differential geometry to
Bayesian estimation of spectral density",
ICIAM 2011, Vancouver, BC, CANADA, July,
2011.

6. 研究組織

(1) 研究代表者

田中 冬彦 (TANAKA FUYUHIKO)
東京大学・大学院情報理工学系研究科・助教
研究者番号：90456161

(2) 研究分担者

()

研究者番号：

(3) 連携研究者

()

研究者番号：