

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25年4月30日現在

機関番号：32670
 研究種目：基盤研究（C）
 研究期間：2009～2012
 課題番号：21500283
 研究課題名（和文）高次元データを解析するための多変量統計モデルと統計的推測手法の理論研究とその応用
 研究課題名（英文）Study on multivariate Statistical models and their inference procedures for analysis of high-dimensional data and its applications
 研究代表者
 今野 良彦（Yoshihiko KONNO）
 日本女子大学・理学部・教授
 研究者番号：00205577

研究成果の概要（和文）：

高次元多変量正規モデルのもとで経験分散共分散行列が特異な場合に、改良型縮小推定量を与えた。特異 Wishart 行列に対する The Stein-Haff identities と calculus on eigenstructures に基づき、2乗リスク関数のもとで経験分散共分散行列を一様に改良することを示した。

研究成果の概要（英文）：

The problem of estimating large covariance matrices of multivariate real normal and complex normal distributions is considered when the dimension of the variables is larger than the number of sample size. The Stein-Haff identities and calculus on eigenstructures for singular Wishart matrices are developed for real and complex cases, respectively. By using these techniques, the unbiased risk estimates for certain class of estimators for the population covariance matrices under an invariant quadratic loss functions are obtained for real and complex cases, respectively.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2009年度	1,100,000	330,000	1,430,000
2010年度	900,000	270,000	1,170,000
2011年度	700,000	210,000	910,000
2012年度	700,000	210,000	910,000
年度			
総計	3,400,000	1,020,000	4,420,000

研究分野：総合領域

科研費の分科・細目：情報学・統計科学

キーワード：ウィシャート行列，縮小型推定量

1. 研究開始当初の背景

統計的多変量解析の出発点は、J. Wishart によるいわゆる Wishart 分布の一般形の導出の成功においてである。この結果は、Siegel

により導入された多変量ガンマ関数へと発展し、数論においても重要な役割を果たすものである。また、Wishart 分布は、大規模ランダム行列理論においても中心的なモデル

のひとつである。多変量解析の出発点である Wishart 分布は、統計的多変量解析の理論において果たした役割はいうに及ばず、数論や理論物理等において分野横断的に研究対象とされる素材であり、この分布の拡張とこの分布を基にした新たな統計的最適理論は、古典的な観点の教科書に展開される内容)を超えて、発展しつつある。

一方、科学技術の急速な発展により、遺伝学研究におけるマイクロアレイデータ、量的形質座位の解析、蛋白・プロテオーム解析、金融工学における株値データ、移動通信や SAR (synthetic aperture radar system) データの解析、顔認証システムにおける画像データのような新しいタイプのデータ (標本数よりもデータの次元数がより高い) に対する統計モデリングと推測手法の理論の本質的な展開 (古典的な枠組みを超えたもの) が必要となってきた。すなわち、標本数よりもデータの次元数がより高い新しいタイプのデータ (以下では、高次元データとよぶことにする) に基づく解析の必要性が高まっている。これらの新しいタイプのデータを解析するためには、

- (1) 高次元データの解析に有効なモデルの考案と検証,
- (2) 高次元データを解析するためのモデルのもとでの新しい手法の提案とその最適性の研究,
- (3) 旧来の手法を用いる場合には、高次元データを解析するためのモデルのもとでの最適性の研究,
- (4) 数値実験による有限標本での精度の検証,
- (5) 応用問題における新たな理論と手法の適応可能性の検証,

を可能な限り体系的に研究することが重要かつ必要である。

これまでの本研究代表者による研究成果において、任意の対称錐 (すなわち、自己双対で等質な開凸錐体で正値対称実行列やエルミート対称複素行列をその例として含むもの) がジョルダン代数によって記述される事実を援用すれば、正値対称実行列の空間上の Wishart 分布に基づく推測理論が任意の対称錐にも見通しよく拡張・適応できることを示した。このことは、多変量実正規分布と複素正規分布それぞれで行われてきた議論はいうまでもなく、Lorentz 錐や Stiefel 多様体上の統計モデルまでの推測理論が統一的な枠組みから議論をすることが期待できることを示唆している。これは、正値対称実行列の空間上に定義される Wishart 分布は多変量分散分析などで基本的な役割を果たし、エルミート対称複素行列の空間上で定義される複素 Wishart 分布は、移動通信や SA このような背景を踏まえて、高次元データの解析のために必要な統計解析手法の開発と最適性の研究を統一的な視点から行うことが肝要であると言えよう。

2. 研究の目的

以上の背景と多変量解析の理論研究の流れが新たな局面へとむかいつつある研究状況、並びにこれまでの本研究代表者の研究成果を踏まえて下記について研究を進めた：

- (1) 高次元データの解析に必要な統計手法とその最適性の理論研究をランダム行列の理論を援用して展開していくこと。とくに、高次元データの設定のもとでの Wishart 分布のスケール行列の推定問題と検定問題の最適性の研究を展開すること
- (2) 因果関係を記述するグラフィカルモデルを利用した高次元データの統計解析手法

を考案し、それらの最適理論を調べること；

- (3) Bayes 予測理論の観点から高次元データの設定のもとでの Wishart 分布に関するよい事前分布を提案し、Bayes 予測に関するリスクを調べること；
- (4) 高次元データのもとでの Wishart 分布の期待値母数の検定問題における尤度比検定統計量等の漸近分布やその精度の向上を目指す近似方式を検討し、古典的多変量解析で知られている結果を統一的かつ体系的に拡張すること；
- (5) 等質錐(自己双対性を仮定しない対称錐を含む広い錐)上の Wishart 分布の期待値母数の推定問題について統計的決定理論の観点から検討し、縮小推定法の構成とその最適性を調べること；
- (6) Wishart 分布は自然母数指数分布族に含まれる。自然母数指数分布族の理論の観点から、等質錐上の新たな統計モデル(等質錐上の楕円分布族や指数 dispersion モデルなど)の構成し、その性質を調べること；
- (7) ジョルダン代数を用いて、対称錐上の統計モデルから Stiefel 多様体 Grassman 多様体上の統計モデルの統一的な記述・表現を検討し、体系的な推測理論を確立すること。

3. 研究の方法

多変量正規モデルにおける標本分布論で中心的な役割を果たす Wishart 分布に関する分布論は表現論における対称空間における微分作用素等の研究と重なる部分が多い。表現論でわかっていることと多変量解析でわかっていることがどのように関連しているかを明らかにするとともに、Jordan 代数に

よるアプローチを用いて多変量統計モデルにおける分布論を展開する研究を担当する。さらに、大規模ランダム行列理論(Random Matrix Theory)を援用しながら高次元データに対する統計推測理論の構築をする。また、連携研究者間の研究指向や背景は非常に異なるので、必要に応じて対面で討論を行いつつ、円滑に研究が進むように連携研究者と研究協力者の研究交流の調整と全般の研究の統括をしながら研究を進めた。

4. 研究成果

- (1) 高次元データの設定のもとでの Wishart 分布のスケール行列の推定問題における最適理論の研究を大きな成果を得た。標本数に変数の次元より小さいときには、素朴な推定量である経験共分散行列は、正則ではなくなり、推定量としての安定性を失う。この欠点を回避するための新たな推定量として、経験共分散行列と単位行列の一次結合の形の推定量を新たに提案したある推定量のクラスのリスク関数の不偏推定量を導出することで、一次結合の係数を適切に求めることができた。さらに、数値実験により、提案した推定量のリスクを具体的に調べた。
- (2) 複素ウィシャート分布のスケール行列の推定問題を不変な 2 乗損失もとで考え、経験共分散行列と単位行列の一次結合の形の推定量の危険関数(推定量の精度を測る指標)の不偏推定量を導出することにより、改良型推定量を求めた。
- (3) 観測値変数と打ち切り変数に独立を仮定した解析手法についての先行研究はいくつかある。しかし、先行研究で提案されているノンパラメトリック手法の妥当性は独立性なしでは保障されない。従属性がある打ち切りデータに対して、母数モデルを仮定することにより、有効な推定手法を提案した。さらに、数値実験により、提案した推定手法の精度を

検証した。この結果については論文発表をした。また、乗数中心極限定理を用いて適合度検定統計量の漸近表示を求めることにより、母数モデルの適合度検定等計量の棄却域を算出するための効率的なアルゴリズムを考案した。

(4) 分散共分散行列が未知の場合の複素行列値多変量正規分布の平均行列の推定問題を不変2乗損失関数のもとで考察した。実数値多変量正規分布の平均行列の推定問題の先行研究結果を参照して、複素行列値多変量正規分布の平均行列の推定量の族である種の変換群のもとで不変なもの考えた。この推定量族のリスクを陽に求めることはできない。そこで、不変推定量族のリスク関数の不偏推定量を導出することにより、リスク関数の評価を可能にし、縮小型推定量を導出のための十分条件を得た。この成果については、チュニジアの研究集会で口頭発表した

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 4 件)

(1) Emura, T. and Konno, Y., A goodness-fit test for parametric models based on dependently truncated data, Computational Statistics and Data Analysis, vol 56(2012), 2237-2250..

(2) Emura, T. and Konno, Y., Multivariate normal distributions approaches for dependently left-truncated data, vol. 53(2012), No. 1, 133-149.

(3) Konno, Y., Estimation of multivariate complex normal covariance under an invariant quadratic loss, Communication

in Statistics -- Theory and Method vol. 39(2010), 1490 - 1497.

(4) Konno, Y., Shrinkage estimators for large covariance matrices in multivariate real and complex normal distributions under an invariant quadratics loss, Journal of Multivariate Analysis, vol 100(2009), 2237-2253.

[学会発表] (計 2 件)

(1) Konno, Y., Shrinkage estimators for large covariance matrices in multivariate normal distribution, The 2nd iims-APRM2012 Meeting (July 2-4, 2012 at Tsukuba) (2012 年 07 月 03 日発表)

(2) Konno, Y., Shrinkage estimation of a mean matrix of a multivariate complex normal distribution, CIMPA-MICINN-UNESCO Research School: Analytical and algebraic tools in Statistics and Graphical Models, September 5 -17, 2011 Hammamet, TUNISIA (2011 年 09 月 07 日発表)

[図書] (計 0 件)

[産業財産権]

○出願状況 (計 0 件)

名称 :

発明者 :

権利者 :

種類 :

番号 :

出願年月日 :

国内外の別 :

○取得状況 (計 0 件)

名称 :

発明者：
権利者：
種類：
番号：
取得年月日：
国内外の別：

[その他]
ホームページ等
<http://mcm-www.jwu.ac.jp/~konno/resarch.html>

6. 研究組織

(1) 研究代表者 今野良彦 (Yoshihiko KONNO)
日本女子大学・理学部・教授
研究者番号：00205577

(2) 研究分担者
()

研究者番号：

(3) 連携研究者
()

研究者番号：