

研究目的

本欄には、研究の全体構想及びその中で本研究の具体的な目的について、冒頭にその要旨を記述した上で、適宜文献を引用しつつ記述し、特に次の点については、焦点を絞り、具体的かつ明確に記述してください。(記述に当たっては、「科学研究費補助金(基盤研究等)における審査及び評価に関する規程」(公募要領 52~99 頁参照)を参考にしてください。)

- ① 研究の学術的背景(本研究に関連する国内・国外の研究動向及び位置づけ、応募者のこれまでの研究成果を踏まえ着想に至った経緯、これまでの研究成果を発展させる場合にはその内容等)
- ② 研究期間内に何をどこまで明らかにしようとするのか
- ③ 当該分野における本研究の学術的な特色・独創的な点及び予想される結果と意義

統計的多変量解析は、相関があると考えられる多変量の観測データを分析する手法の理論の体系である。科学技術の急速な発展から新しいタイプのデータを解析(たとえば、遺伝学研究におけるマイクロアレイデータと量的形質座位の解析等が典型的な例)するために、古典的な多変量解析の理論的な枠組み(Andersonの教科書[1]を参照)を超えて、モデリングと統計推測手法の研究を展開する必要がある。本研究では、分野横断的な数理理論を援用することにより、統計的多変量解析の古典的な理論を超えた体系的かつ一般的な推測理論を構築することを目的とする。さらに、遺伝学研究におけるマイクロアレイデータと量的形質座位の解析への応用を視野にいれて研究する。

① 研究の背景について: 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(Anderson[1]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、

- (α) 高次元データの解析に有効なモデルの考案と検証,
- (β) 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,
- (γ) 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,
- (δ) 数値実験による有限標本での精度の検証,
- (ε) 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 [2] 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実行列をその例として含むもの)がジョルダン代数(すなわち、可換律のほかにジョルダン積という特殊な結合律をみたす代数系)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

研究目的 (つづき)

をすることが期待できることを示唆している．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 何をどこまであきらかにしようとするか: 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ていく：

- (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 多様体と 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 学術的な特色・独創的な点および予想される結果とその意義 : (あ) 本研究の学術的な特色は，研究対象と手法は，ランダム行列の理論，調和解析学，線形計画法，情報理論，通信工学，統計的多変量解析等の広い分野で縦断し学際的なことである；(い) 古典的な多変量解析を新たな視点から大幅に発展させるものである；(う) 等質錐上の多変量解析は世界的にも注目が集まっている最新の話題である；(え) 本研究の目標とする結果は，今後大きく発展すると予想される等質錐上の多変量解析の進展と建設に先鋭的かつ本質的な貢献になることが期待できる；(お) 本研究の結果が広い分野に応用できる可能性も秘めている．また，機械学習理論，ネットワーク理論，遺伝子モデル，計量経済学モデル等との幅広い分野とも非常に関係が深く，これらの分野への新しい多変量モデルの応用等その実用的な面においても研究の意義はおおきい．

参考文献

- [1] T.W. Anderson, An Introduction to Multivariate Statistical Analysis, 3rd. ed., Wiley (2003).
- [2] Y. Konno, Estimation of a normal covariance matrix parametrized by irreducible symmetric cones under Stein's loss, *J. Multivariate Anal.* **98** (2007) 295–316.