

| | |
|--------------|---|
| Title | Statistical Inference Concerning a High-Dimensional Mean Vector |
| Author(s) | 片山, 翔太 |
| Citation | 大阪大学, 2013, 博士論文 |
| Version Type | |
| URL | https://hdl.handle.net/11094/59859 |
| rights | |
| Note | 著者からインターネット公開の許諾が得られていないため、論文の要旨のみを公開しています。全文のご利用をご希望の場合は、 〈a href="https://www.library.osaka-u.ac.jp/thesis/#closed"〉 大阪大学の博士論文について 〈/a〉 をご参照ください。 |

Osaka University Knowledge Archive : OUKA

<https://ir.library.osaka-u.ac.jp/>

Osaka University

【62】

| | |
|------------|---|
| 氏名 | かた やま しょう た 片 山 翔 太 |
| 博士の専攻分野の名称 | 博 士 (理学) |
| 学位記番号 | 第 2 6 1 1 0 号 |
| 学位授与年月日 | 平成 25 年 3 月 25 日 |
| 学位授与の要件 | 学位規則第 4 条第 1 項該当 基礎工学研究科システム創成専攻 |
| 学位論文名 | Statistical Inference Concerning a High-Dimensional Mean Vector (高次元平均ベクトルに対する統計的推測) |
| 論文審査委員 | (主査) 教 授 狩野 裕 (副査) 教 授 内田 雅之 教 授 下平 英寿 |

論 文 内 容 の 要 旨

近年、高次元データ、すなわち変数の次元が標本サイズよりもはるかに大きいもしくは同程度のデータが様々な統計的応用の場面において観測されている。しかしながら、従来の多変量解析手法の多くは低次元大標本の枠組みで導出されているため、高次元データに適用することができない。そのため、高次元データに対して適切な手法を開発することは重要である。本論文では特に、高次元データにおける平均ベクトルの検定問題及び、推定問題を扱う。

平均ベクトルの検定問題に関しては、高次元データにも適用可能な検定統計量が現在までに幾つか提案されており、高次元漸近枠組み（変数の次元と標本サイズが共に大きくなる）の下で、それらの漸近分布も導出されている。本論文では先ず、既存の検定統計量の漸近帰無分布が母共分散行列の構造に依存することを示し、既存手法では実際の高次元データ解析の際に問題が生じることを指摘する。すなわち、既存手法を用いる場合、母共分散行列の構造を誤推

定してしまうと、誤った検定結果を返してしまう。これを解決するために、本論文では次に、高次元漸近枠組みにおける漸近帰無分布が母共分散行列の構造に依存しないような検定統計量を提案する。その一方で、高次元データにおけるBehrens-Fisher問題において、標本の尺度変換に関して不変な検定統計量を提案し、その高次元漸近枠組みにおける漸近分布も導出する。

平均ベクトルの推定問題において、伝統的な推定量は標本平均ベクトルである。しかしながら、高次元データにおいて、標本平均ベクトルは母平均ベクトルをL2ノルムの意味で一致推定することができない。高次元データに対して適切な推定量を導くため、本論文では、母平均ベクトルがスパース（多くの要素が0）であるという制約を課し、その制約条件の下で正規尤度関数を最大化する推定量を提案する。高次元漸近枠組みの下、提案する推定量はスパース構造を完全に復元できることを示し、その平均二乗誤差も導出する。

論文審査の結果の要旨

近年、ビッグデータ解析と呼ばれる(超)高次元データの分析方法の開発が喫緊の課題である。高次元データとは変数の次元が標本サイズよりもはるかに大きいもしくは同程度のデータを言う。分析手法に関しては、従来の多変量解析手法は低次元大標本の枠組みで導出されているため、高次元データのそれとは全く別物と考えた方がよい。申請者は、本論文において高次元データにおける平均ベクトルの統計的推測法を発展させた。

平均ベクトルの検定問題に関しては、高次元データにも適用可能な検定統計量が現在までに幾つか提案されており、高次元漸近枠組みの下で、それらの漸近分布も導出されている。本論文ではまず、既存の検定統計量の漸近帰無分布が母共分散行列の構造に依存し、固有値のあり様によって、漸近分布が全く異なることを示した。すなわち、既存手法を用いる場合、母共分散行列の構造を誤特定してしまうと、誤った検定結果となり得る。この問題を解決するために、本論文では、高次元漸近枠組みにおける漸近帰無分布が母共分散行列の構造に依存しないような検定統計量を提案した。また、高次元データにおけるBehrens-Fisher問題において、標本の尺度変換に関して不変な検定統計量を提案し、その高次元漸近枠組みにおける漸近分布を導出した。

平均ベクトルの推定問題において、伝統的な推定量は標本平均ベクトルである。しかしながら、高次元データにおいて、標本平均ベクトルは母平均ベクトルをL2ノルムの意味で一致推定することができない。高次元データに対して適切な推定量を導くため、本論文では、母平均ベクトルがスパース（多くの要素が0）であるという仮定を導入し、この下で正規尤度関数を最大化する推定量を提案した。高次元漸近枠組みの下、提案する推定量はスパース構造を完全に復元できることを示し、さらに、その平均二乗誤差を導出した。

申請者による「高次元データ解析の理論的発展」は統計学への有意な貢献であり、博士（理学）の学位論文として価値のあるものと認める。