

## 審査の結果の要旨

論文提出者 田中研太郎

近年のデータ観測技術の発達により、大規模なデータが容易に得られるようになってきた。これにともない、データマイニング、パターン認識、機械学習などの大規模データの統計的解析手法の需要が高まっている。これらの統計的解析において、混合正規分布モデルと呼ばれる確率モデルが広く用いられている。混合正規分布モデルは、正規分布の密度関数を重みを付けて足しあわせたものであり、パラメータを調整することにより、多峰性を持つ確率密度関数など、複雑な確率分布を構成することが可能である。一般的に、確率モデルのパラメータは最尤推定法によって推定することが多い。しかしながら、混合正規分布モデルのパラメータの最尤推定には、推定量の一致性の欠如という重大な問題点があることが知られている。

本論文は“Strong consistency of maximum likelihood type estimators for finite mixture models (有限混合分布モデルに対する最尤型推定量の強一致性)”と題し、混合分布モデルにおける問題点を解決する方法として、制約付き最尤推定量と罰則付き最尤推定量を提案し、その推定量の理論的な性質について解明している。本論文では正規分布のみならず、一般の位置尺度分布族の有限混合分布についての結果を導出している。本論文は、概要から結論の章までを含め全部で5章よりなる。

第1章“Introduction”では混合分布モデルに関するいくつかの問題点を挙げ、本論文で扱う課題の位置づけを述べている。また、論文全体の流れについて説明している。

第2章“Finite Mixture Models”においては、まず、本論文で扱う混合位置尺度分布モデルについて定式化している。そして、混合位置尺度分布モデルにおいて各成分の位置母数と尺度母数の値が未知である場合に、尺度母数を0に近づける場合を考えると、尤度関数が非有界となることを示し、最尤推定が実行できないという問題点があることを説明している。また、sieve法やベイズ法など、関連する先行研究について概説している。

さらに、最尤推定に基づいた通常のEMアルゴリズムを用いてパラメータを推定しようとする、尤度関数の非有界性により、アルゴリズムが破綻してしまうことを数値実験により示している。

第3章“Constrained Maximum Likelihood Estimator”では、混合分布の各成分の尺度母数の値を下から押さえるという制約を付けた制約付き最尤推定量を提案している。これにより、尤度関数の非有界性に起因する問題点は回避できる。また、制約は標本サイズの増加とともに緩められることを想定しており、制約を

指数関数的なあるオーダー以下で緩めていったときに、制約付き最尤推定量が強一貫性を持つことを証明している。さらに、強一貫性はもっと速いオーダーで制約を緩和すると破綻してしまうことも証明している。そして、制約付き最尤推定量を求めるための制約付き EM アルゴリズムを提案し、いくつかの数値実験でその有効性を示している。クロスバリデーションによって制約の強さの値を選択する方法についても論じている。

第 4 章 “Penalized Maximum Likelihood Estimator” では、2 つの種類の罰則付き最尤推定量を提案している。1 つめは、混合分布の各成分の尺度母数の値そのものに対して罰則を付ける方法である。尺度母数が 0 に近づくとともに、それに反発するような罰則を付けることにより、尤度関数の非有界性の問題を回避している。そして、罰則が十分強ければ、罰則付き最尤推定量が強一貫性を持つことを証明している。また、この罰則付き最尤推定量を求めるための罰則付き EM アルゴリズムについても考察し、適切な罰則を用いることで、良い推定値が得られることを数値実験で示している。2 つめは、混合分布の各成分の尺度母数同士の比に対して罰則を付ける方法である。尺度母数の比が 0 に近づくと反発するような罰則を付けた罰則付き最尤推定量を考えると、それが強一貫性を持つことを証明している。また、これら 2 つの種類の罰則は、標本サイズの増加とともに一貫性を保ったまま緩和可能であることも証明している。その結果を使って、尺度母数の比に標本サイズに依存する制約を付けた場合の制約付き最尤推定量が強一貫性を持つことも証明している。この制約付き最尤推定量の強一貫性については、長いこと混合分布における未解決問題として知られており、本論文の結果により肯定的に解決された。

第 5 章 “Conclusion” では本論文の結果のまとめを与えている。

以上を総合するに、本論文は、統計科学において広く使われる混合分布モデルの推定における重大な問題点に対し、新たな視点から問題の解決方法を提供し、また、その理論的な性質を解明しており、数理情報学の分野の発展に大きく寄与するものである。

よって本論文は、博士（情報理工学）の学位請求論文として合格と認められる。