

氏 名 吉田 亮

学位（専攻分野） 博士（学術）

学位記番号 総研大甲第 803 号

学位授与の日付 平成 16 年 9 月 30 日

学位授与の要件 複合科学研究科 統計科学専攻  
学位規則第 6 条第 1 項該当

学位論文題目 Mixture Models in Bayesian Computation and Mixed  
Factors Analysis

論文審査委員 主 査 教授 江口 真透  
助教授 南 美穂子  
助教授 藤澤 洋徳  
教授 石井 信（奈良先端科学技術大学院  
大学）

## 論文内容の要旨

提出された論文は全6章120頁からなり、以下の構成である。

第1章は、本論文の概要を与えた章で、統計科学の様々な応用領域における混合モデルの重要性について言及している。本研究の主目的である「混合モデルによるベイジアンコンピュータショナルアルゴリズムの開発」と「遺伝子発現パターンの統計的モデリング」に関する研究背景を概観しながら、本研究の学問的位置付けおよび応用面における重要性について述べられている。

第2章は、混合モデルの定義、尤度原理に基づく推定方法について述べ、その問題点を指摘している。また、近年、提案されてきたいくつかの混合モデルの研究事例について、数理的および応用的観点からレビューを行っている。

第3章は、ベイズ計算のためのアルゴリズムに関する研究である。ベイズ推定の多くの局面で直面する数値計算の困難さについて言及し、その問題に対するひとつの克服手段である Importance Sampling 法 (以下 IS 法) と Sampling Importance Resampling (以下 SIR) アルゴリズムについて紹介している。IS 法と SIR アルゴリズムにおけるシミュレーション効率性が、個々の問題によって定義される最適サンプリング分布と実際に用いるサンプリング分布との距離に依存することを示した上で、Parametric Adaptive Importance Sampling (以下 PAIS) というアルゴリズムを提案している。PAIS ではサンプリング分布の族として、混合モデルなどの適当なパラメトリックモデルを仮定し、その中で最適分布に最も近いメンバーが探索され、ひとつのサンプリング分布が自動的にあたえられる。

第4章では、超高次元データに対するクラスタリングの新しい分析手法「混合因子分析法」を提案している。マイクロアレイデータのクラスター分析における特徴データの特殊性について言及し、既存の統計的方法では困難とされる問題を克服するために、因子分析の拡張を行っている。この方法の特筆すべき点として、(1) マイクロアレイデータのような超高次元特徴データに対しても、そのモデル推定が可能なこと、(2) さらに推定に要する計算が容易であること、(3) 推定された因子と元々の特徴ベクトルとの依存性(相関)をみることで、相互作用によって機能し、かつ分類に寄与する変数群の特定が可能となること、が主張されている。提案する推定方法の計算アルゴリズムによる反復推定が逐次的な主成分分析、逐次的なフィッシャー線形判別法と密接に関係していることを明らかにしている。

第5章は、遺伝子発現データへの混合因子分析の適用結果が示されている。混合因子分析を実際のマイクロアレイデータに適用し、その遺伝子発現解析における有効性を示している。

第6章はまとめである。

以上のように本申請論文は IS 法とクラスタリングの内容で2つの提案を混合モデルの統一的な観点から行っている。特に、提案された「混合因子分析法」は、マイクロアレイデータの解析に著しい成果が見られ、従来の幾つかの困難な点が克服されている。

## 論文の審査結果の要旨

本申請論文は、IS法とクラスタリングの内容で2つの提案を混合モデルの統一的な観点から行っている。特に、提案された「混合因子分析法」は、マイクロアレイデータの解析に著しい成果が見られ、従来の幾つかの困難な点が克服されている。これらの結果は、バイオインフォマティクスの国際会議でも発表され（査読付き国際会議 Proceeding）、また審査付き学術雑誌に一編、採択されている。

このようなことから本申請論文が学位授与に十分に値するという結論に達した。