

審査の結果の要旨

氏名 二見 太

自然科学，社会科学，情報科学の様々な分野において，観測データの生成メカニズムの分析に確率モデルに基づくベイズ推論がよく用いられる．多くの現実的な確率モデルに対して，ベイズ推論は解析的に実行できないため，これまで様々な近似法が研究されてきた．しかし，従来の近似法は，計算効率が良くない場合や，観測データの外れ値に対してロバストでない場合があるという問題があった．このような背景のもと，本博士論文では，これらの問題を緩和する新しい近似ベイズ推論手法を提案している．本論文は以下のように英文6章からなっている．

1章「序論」では，機械学習やベイズ推論の背景を説明するとともに，本論文で提案する手法の概要および本論文の構成について述べている．

2章「準備と関連研究」では，まず本論文の主題であるベイズ推論とロバスト推論に関する基礎概念を導入している．そして，ベイズ推論で用いられる代表的な近似手法やロバスト推論で広く用いられる影響関数を紹介し，どのような場合にロバスト性を得られるのかについて論じている．

3章「 t 指数分布族とそれに対する期待値伝播法」では，ベイズ推論の枠組みの中で外れ値に対するロバスト性を得る手段の一つとして，確率モデル自体を修正するアプローチを議論している．具体的には，モデルに含まれるガウス分布を，より裾が重いスチューデント t 分布に置き換えることによりロバスト性を得る方法を扱っている．しかし，スチューデント t 分布はガウス分布のように指数分布族に属さないため，計算効率の高い近似アルゴリズムを開発することが困難である．本章では， q -代数とよばれる特殊な代数系を用いれば，スチューデント t 分布に対しても指数分布族と同等な解析的計算が可能であることを利用し，指数分布族に対する期待値伝播法と呼ばれる近似手法を，スチューデント t 分布を含む一般化された指数分布族（ t 指数分布族）に対して適用できるように拡張した．そして，この拡張した近似手法に基づいて，新しい分類モデルであるスチューデント t 過程分類器を開発し，そのロバスト性を計算機実験により示した．

4章「ロバストな距離尺度を用いた変分推論」では，外れ値に対するロバスト性を得るもう一つの手段である，学習法そのものを変更するアプローチについて論じている．3章で示したモデルを修正するアプローチでは，機械学習で近年よく用いられるニューラルネットワークのような複雑なモデルをロバスト化することは難しい．本章では，ベイズ推論はある距離尺度のもとでの最適化問題の解と捉えられるという事実に基づいて，最適化問題にロバストな距離尺度を導入することによって，外れ値の影響を系統的に抑制できる推論手法を提案した．更に，変分推論を用いることによって計算効率の良いアルゴリズムを構成し，影響関数を用いてそのロバスト性を理論的に示した．

5章「フランク・ウルフのアルゴリズムを用いた逐次近似ベイズ推論」では，ベイズ推論の近似アプローチの一つであるサンプリング手法の計算効率について論じている．3，4章で論

じた近似手法はパラメトリックなアプローチであり、計算効率が良いという長所があるが、パラメトリックな仮定に起因するバイアスが問題となり得る。一方、サンプリングに基づく近似では、多数の標本を用いることによって任意の分布を十分な精度で近似できる事が理論的に保証される。しかし、高次元や多峰性を持つ複雑な分布に対しては、計算量が莫大になってしまうという短所がある。そこで本章では、それぞれの手法の長所を組み合わせることのできる新しい近似手法を提案している。具体的には、最大平均距離と呼ばれる距離尺度を再生核ヒルベルト空間上で最小化するという制約付き凸最適化問題を構成し、それをフランク・ウルフのアルゴリズムによって逐次的に解くというアルゴリズムであり、計算効率が良く、近似性能が理論的に保証されるという優れた性質を有する。

6章「まとめと今後の展望」では、結論と今後の課題について述べている。

以上を要するに、本論文は、機械学習の標準的な手法である近似ベイズ推論の計算効率および観測データの外れ値に対するロバスト性を大幅に改善できる革新的な手法を提案するものであり、開発されたアルゴリズムは、有用性が理論的に保証されるとともに、実際的な有用性も計算機実験により示されている。このような研究成果は、複雑理工学、特に複雑系科学の理解のために必要な機械学習分野に対する貢献が大きい。

よって本論文は博士（科学）の学位請求論文として合格と認められる。

以上1879字