

氏 名 金川 元信

学位(専攻分野) 博士(統計科学)

学位記番号 総研大甲第 1829 号

学位授与の日付 平成28年3月24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 Empirical representations of probability distributions via
kernel mean embeddings

論文審査委員 主 査 教授 栗木 哲
教授 福水 健次
准教授 持橋 大地
准教授 吉田 亮
准教授 鈴木 大慈 東京工業大学大学院

論文内容の要旨
Summary of thesis contents

博士申請論文は、正定値カーネルを用いたカーネル平均埋め込み (kernel mean embedding) と呼ばれる積分変換による、確率分布のサンプル表現 (empirical representation) に関して論じたものである。英文で書かれており、全 6 章、108 頁からなる。

第 1 章は、本論文の序章である。まず研究の背景として、カーネル平均埋め込みが統計的推論の枠組みを踏まえてサーベイされている。確率分布を表現するために通常行われる方法として、パラメトリックモデルによる表現、特性関数による表現があるが、それらと対比する形でカーネル平均埋め込みが導入されている。また経験分布、モンテカルロ法の説明の後、本論文の学問的貢献として (i)カーネル平均埋め込みとサンプリングの融合法の提案、(ii) (i)に基づく新しいフィルタリング法であるカーネルモンテカルロフィルタの提案、(iii)カーネル平均埋め込みの推定量によるさまざまな統計量の推定に関する理論解析、の 3 点が示されている。

また第 2 章では、本論文で必要な定義と概念の準備が行われている。正定値カーネルと再生核ヒルベルト空間 (reproducing kernel Hilbert space, RKHS)、分布の特徴づけを可能とするカーネル、条件付きカーネル平均、カーネルベイズルール、重み付きサンプリング、カーネルハーディング法に関する基礎事項が整理されている。(ハーディング法とは、逐次的最適化の一種であり、統計学における逐次最適実験計画と同種のものである。) とくにカーネル平均の経験表現を導入した後に、その推定量が負の重みを許す重み付きサンプルによって表現されることが説明されている。

第 3 章から第 5 章において、本論文の主要な結果が述べられている。まず第 3 章では、カーネル平均埋め込みによる分布推定における、サンプリングとリサンプリング法を提案している。まず、ある分布 (対応する確率変数を X とする) と条件付き確率 ($p(y|x)$ で表される) から新しい分布 (確率変数を Y とする) を得る和法則をカーネル平均埋め込みによって実現する方法として、条件付き確率からのサンプリングを用いる方法を提案しその推定量の収束性を議論している。ここで X の分布はサンプル、すなわち経験分布で表現され、 Y の分布もサンプルで表現することを前提とする。最初に、収束のレートは X のサンプル表現の精度と、サンプリングの重みの関数として陽に表現されることが示されている。この結果に基づいて、サンプリングによる推定量の精度改善を目的として、カーネルハーディングを用いたリサンプリング法が提案され収束レートの改善が理論的に示されている。また計算機実験によってその有効性が検証されている。

第 4 章は、カーネル平均埋め込みを用いたカーネルモンテカルロフィルタを提案している。観測モデルが未知であるが状態と観測の組の訓練データが得られる状況を想定し、第 3 章の結果を応用して、カーネル平均埋め込みによる状態変数のベイズ推定法が議論され、サンプリング、カーネルベイズ則、リサンプリングを用いたフィルタリング法を提案している。この方法の収束性を理論的に示すとともに、この方法の優位性を示すために、画像情報によって建物内を循環するロボットの位置推定問題に応用し従来手法との比較を行っている。比較は、 k -最近接法、ガウス過程回帰法といった手法との対比で行われ、提案法の優位性を示している。

(別紙様式 2)
(Separate Form 2)

第 5 章はカーネル平均埋め込みの推定量として重み付きサンプル表現が与えられた場合に，関数の期待値を重み付き平均値で推定する方法に関して論じている．最初に関数空間 (Sobolev 空間および Besov 空間)に関する準備がなされ，平均埋め込みに用いる正定値カーネルが定める再生核ヒルベルト空間よりも広いこれらの関数空間に属する関数に対して，期待値の収束レートを理論的に導いている．さらに重み付きサンプルと平滑化カーネルの畳み込みによる確率密度推定法を提案し，収束レートを示している．本章の内容は高度に数学的であるが，理論を確認するための数値計算が大規模になされている．

第 6 章は論文のまとめである．将来への研究課題として，カーネル法と MCMC の結合，リサンプリング法の加速，カーネルモンテカルロ法に基づくパラメータ推測などが挙げられている．参考文献リストは 8 頁 (論文数 94) にわたる．

なお，第 3 章，第 4 章の内容は，査読付き学術雑誌

Kanagawa, M., Nishiyama, Y., Gretton, A., and Fukumizu, K. (2016). Filtering with State-Observation Examples via Kernel Monte Carlo Filter, *Neural Computation*, volume 28, issue 2, pages 382-444

および，査読付き国際学会プロシーディングス

Kanagawa, M., Nishiyama, Y., Gretton, A., and Fukumizu, K. (2014). Monte Carlo filtering using kernel embedding of distributions, in *Proceedings of the 28th AAAI Conference on Artificial Intelligence (AAAI-14)*, pages 1897-1903

として，また第 5 章の内容は査読付き国際学会プロシーディングス

Kanagawa, M. and Fukumizu, K. (2014). Recovering distributions from Gaussian RKHS embeddings, in *Proceedings of the 17th International Conference on Artificial Intelligence and Statistics (AISTATS 2014)*, pages 457-465

として出版済みである．

(別紙様式 3)
(Separate Form 3)

博士論文の審査結果の要旨

Summary of the results of the doctoral thesis screening

カーネル平均埋め込みは、新しいノンパラメトリック推定の方法として機械学習分野で近年発展してきた。3章で提案されている方法は、カーネル平均埋め込みとサンプリングを組み合わせしており、ノンパラメトリックなカーネル法と既知の確率モデルを組み合わせ推論するという新しい方向性を打ち出した点で意義が大きい。またカーネルハーディングを用いたサンプリング法の新規性は高く、4章ではこれらを応用することによって、実用的に有望なフィルタリング法を提案した意義は大きい。さらに5章で論じている広い関数クラスに対する期待値の収束性は、従来理論的に未解決であった問題に一定の解答を与えたもので、平均、分散、信頼区間など重要な統計量の推定法に理論的根拠を与えた点で価値が高い。なお3章および4章の内容は査読付き国際学術雑誌 *Neural Computation* に、5章の内容は査読付き国際会議 *International Conference on Artificial Intelligence and Statistics (AISTATS 2014)* 予稿集に採録されている。

以上の状況に基づき、本審査委員会は、申請者の学位請求論文が学位に十分に値する水準にあると全員一致で判定した。