

論文 / 著書情報
Article / Book Information

題目(和文)	頑健かつオンライン学習可能なノンパラメトリック密度推定法
Title(English)	
著者(和文)	中村圭宏
Author(English)	Yoshihiro Nakamura
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第10230号, 授与年月日:2016年3月26日, 学位の種別:課程博士, 審査員:長谷川 修,渡邊 澄夫,樺島 祥介,新田 克己,小野 功
Citation(English)	Degree:., Conferring organization: Tokyo Institute of Technology, Report number:甲第10230号, Conferred date:2016/3/26, Degree Type:Course doctor, Examiner:,,,,,
学位種別(和文)	博士論文
Category(English)	Doctoral Thesis
種別(和文)	審査の要旨
Type(English)	Exam Summary

論文審査の要旨及び審査員

報告番号	甲第	号	学位申請者氏名	中村 圭宏		
論文審査 審査員		氏名	職名		氏名	職名
	主査	長谷川 修	准教授	審査員	小野 功	准教授
	審査員	新田 克己	教授			
		渡邊 澄夫	教授			
樺島 祥介		教授				

論文審査の要旨 (2000 字程度)

本論文は、「頑健かつオンライン学習可能なノンパラメトリック密度推定法」と題し、和文 6 章から成っている。

第 1 章「序論」では、近年社会のあらゆる場面で膨大なデータの生成と蓄積が進んでおり、それら进行处理するために、(1) 高速オンライン学習、(2) ノンパラメトリック、(3) 高いロバスト性、の機能を有するデータ駆動型の確率密度分布推定手法を確立することが、極めて重要であると指摘している。また、本研究の目的は、それを実現するアルゴリズムを理論的考察に基づき構築し、有効性を数値実験により確認することにあると述べている。

第 2 章「カーネル密度推定法」では、ノンパラメトリック確率密度推定法であるカーネル密度推定法の従来研究をまとめている。特に、オンライン学習性や、ロバスト性を有する既存のカーネル密度推定法については理論的に詳細に考察し、本研究で取り組むべき課題に言及している。

第 3 章「自己増殖型ニューラルネットワーク」では、本研究の提案手法で利用した、自己増殖型ニューラルネットワーク (Self Organizing Incremental Neural Network、以下 SOINN) について、その学習原理と、確率密度推定の目的に応用可能な統計的な性質について詳細に考察している。

第 4 章「自己増殖型ニューラルネットワークに基づくノンパラメトリック密度推定法」では、本研究で提案する、SOINN の統計的な性質を応用したノンパラメトリック密度推定法の詳細を述べている。具体的には、SOINN の学習の過程で生成されたニューロンに対してそれぞれガウスクERNELを適用し、学習データの確率密度分布の推定は、それらのカーネルの合成関数を得ることで行うと述べている。ここで、ガウスクERNELは、各ニューロンが学習の過程で勝者ニューロンとなった回数に基づき、その形状をオンラインで調整・更新するとしている。さらに提案手法は、SOINN が持つノイズ除去機能を活用することで、学習データに含まれるノイズに対するロバスト性も有すると述べている。この提案手法の有効性・優位性は、UCI Machine Learning Repository の複数のデータセットを用いた数値実験により既存手法と比較することで、定量的に確認している。

第 5 章「学習アルゴリズムの改良による性能向上」では、第 4 章で述べた提案手法の拡張について述べている。既存の SOINN では、閾値領域という制約を導入することで、Competitive Hebbian Learning 法 (以下、CHL) にロバスト性を与えているが、各ノード周辺のサンプルは等方的ガウス分布に従うと仮定しており、勝者ノードの閾値領域内に入った入力サンプルのみが勝者ノードのサンプルとしてカウントされる。このため、既存の SOINN では閾値領域の形状は超球であり、各ノード周辺のサンプルの分布は共分散行列が単位行列の定数倍で、分布の広がり方が方向に依存しないガウス分布になる。これは、本来 CHL が形成するネットワーク構造と矛盾し、CHL では入力サンプルが多様体に沿って分布する場合、多様体に沿った部分にのみエッジが形成され、入力サンプルはネットワークに沿って分布し、あるノードの周辺のサンプルは、そのノードから出ているエッジの方向に広がり、全ての方向に均一に広がってはいないと指摘している。そこで本研究では、閾値領域を局所ネットワーク共分散行列が規定するガウス分布の形状になるよう拡張したほか、高次元空間やサンプル数が少ない場合に十分なエッジが作成されない問題の克服のため、k 近傍グラフに基いてエッジを追加する拡張も加えている。これらの拡張の有効性・優位性も、複数のデータセットを用いた数値実験により既存手法と比較し、定量的に確認している。

第 6 章「結論」では、本論文を総括するとともに、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、データ駆動型の新たな確率密度分布推定手法を提案するとともに、その有効性・優位性を実験的に確認しており、工学上の貢献が大きい。よって本論文は、博士(工学)の学位論文として十分な価値があるものと認められる。