

## 科学研究費助成事業 研究成果報告書

平成 29 年 6 月 7 日現在

機関番号：32665

研究種目：基盤研究(C) (一般)

研究期間：2011～2016

課題番号：23540224

研究課題名(和文) 学習理論における特異性と複素および実解析的考察

研究課題名(英文) Consideration on singularities in learning theory by complex and real analysis

研究代表者

青柳 美輝 (AOYAGI, Miki)

日本大学・理工学部・准教授

研究者番号：90338434

交付決定額(研究期間全体)：(直接経費) 3,800,000円

研究成果の概要(和文)：学習理論の目的は、真の分布から発生する多量のデータセットから、そのデータを発している情報源の真の分布を再生・推測することである。WAIC法はAIC法を一般化したもので、特異学習モデルにも適用可能である。ベイズ推測における学習係数は、特異学習モデルの学習効率をあたえるものでWAIC法の重要な役割を果たしている。学習係数は、代数幾何の分野では、相対エントロピーのlog canonical thresholdとして知られている。この研究では、学習係数に関する考察およびそれらを得るために必要な定理を求めた。

研究成果の概要(英文)：The purpose of a learning system is to estimate an unknown true density function that generates the data. Recently, the WAIC model selection method, which is a generalized AIC, has been considered for estimating a probability function from data in a singular learning system. The learning coefficient in Bayesian estimation serves to measure the learning efficiency, and has an important role in the WAIC method. Mathematically, the learning coefficient is the log canonical threshold of the relative entropy. In this study we consider learning coefficients and several theorems for our results.

研究分野：学習理論

キーワード：学習係数 特異点解消 log canonical threshold

## 1. 研究開始当初の背景

機械学習の研究には、2つのアプローチの方法がある。一つは、画像、音声、遺伝子など、データそのものの特徴に着目して、機械学習システムを構築していく方法、もう一つは、普遍的に従う法則を発見し、その性質について研究する方法である。

第2の方法で、学習システムの挙動を定めている一般的な法則を解明し、その上で情報科学的な設計法を与える理論を学習理論と呼ぶ。この研究は第2の方法で行うものである。

学習理論においては、フィッシャー情報量が2次近似可能な正則モデルと、フィッシャー情報量が2次近似できない、すなわち退化する特異モデルがある。比較的好く用いられるモデル、混合正規分布、三層ニューラルネットワーク、混合二項分布などは、特異モデルである。これらのモデルには、正則モデルに用いられる方法、例えば、モデル選択法であるAIC, TIC, HQ, NIC, BIC, MDLなどは適用できない。このように、古典的な理論の枠組みの中では捉えることができず、近年急速に多くの理論の研究が始まった。

学習理論における学習効率を表す学習係数は、カルバック情報量の実数体上のlog canonical thresholdとその位数で与えられる。log canonical thresholdは対応するゼータ関数の最大の極であり、代数解析におけるBernstein-Sato多項式の最大の根であることが知られている。代数幾何・代数解析では主に代数閉体上での研究が行われている。また、低次元での研究が主である。従って、そのまま複素体上の定理を学習係数に適用することができない。例えば、複素数体上のlog canonical thresholdは1より小さいが、実数体上ではそうとは限らない。学習理論における情報量のlog canonical thresholdは、ほとんどが1より大きい。ある関数族に関しては、実数体上のlog canonical thresholdの方が多くの情報を持っていることが知られている。また、複素数体において、関数を超平面に制限したとき、log canonical threshold値は小さくなるという定理が証明されているが、実数体上では反例がある。このように学習係数を求めることは、数学的観点からも興味のある問題である。

応用上、学習係数は、ハイパーパラメータの設計において、周辺尤度の値の理論値を与えるため、MCMC法の精度を解明することや、精度の改良法、モデル選択解析に応用されている。

## 2. 研究の目的

この研究の目的としては、カルバック情報量の特異性がどのような性質を持つのかということや、その性質と学習係数との関係を明らかにすることであった。特に統計解析で比

較的良好に使用される、混合正規分布、三層ニューラルネットワーク、混合二項分布などの学習モデルから得られるVandermonde matrix型特異点やボルツマンマシンのベイズ推測に関する学習効率について考察することが目的であった。Vandermonde matrix型特異点は、異なるモデルから共通の特異性が現れたことから、学習理論において本質的ではないかと考えられている。また、これらの理論値が得られれば、それを用いて、最適なモデル選択法への応用、効率のよい機械学習の構成法について考察することを計画していた。

## 3. 研究の方法

WAIC法では、BayesianおよびGibbs generalization errorが、真の分布の情報を用いずに、BayesianおよびGibbs training errorから推測できることを示している。training errorは、観測データおよび学習モデルを用いて求められる。実際の応用や実験では、通常、真の分布が不明で、観測データのみが得られている。したがって観測データからの真の分布の推定やモデル選択にWAIC法は有効である。

WAIC法では、学習係数が重要な役割を果たす。log canonical thresholdは、広中の特異点解消定理により、原理的には有限の手続きにより求められるが、具体的に求めるのは難しいとされている。計算機で代数計算により行う方式も提案されているが、特異モデルのカルバック情報量は中間ユニット数などのパラメータを含んでいるため、確定された多項式の特異点解消よりも高度な面を含んでいる。更なる困難な問題点として、特異点が孤立していない・ニュートン図形が退化している等があげられる。

今まで、特異モデルの学習係数は、数例においてその上限が得られていたにすぎず情報科学において長い間未解明であったが、近年、我々により、縮小ランクモデルの場合や、出力層が1個である三層ニューラルネットワークの場合、1次元の混合正規分布の場合の学習係数について解明された。また、帰納的に行う特異点解消の中心となる多様体の適当な選択によって、Vandermonde matrix型特異点の学習係数のバウンドが得られている。

これらの結果を足がかりに、高次元の特異点の研究を行い、その理論結果を元に新たな効率のよい機械学習を構築する方法を模索する。

## 4. 研究成果

23年度：Vandermonde matrix型特異点のlog canonical thresholdの研究を行った。帰納的に行う特異点解消の中心となる多様体の適当な選択の改良によって新たなバウンド

を得た。以前得られた結果より、条件によっては、かなり真値に近づいた。この結果は、論文 "Learning coefficient of generalization error in Bayesian estimation and Vandermonde matrix type singularity, Neural Computation" に掲載されることになった。また、平成 23 年 12 月アメリカ Palo Alto 市 AIM Research Conference Center で開かれた国際会議 "Singular learning theory: connecting algebraic geometry and model selection in statistics" にて、この結果について講演した。

24 年度：Vandermonde matrix 型特異点と、Restricted Boltzmann Machine などに現れる対称型特異点の log canonical threshold の研究を行った。イデアルの包含関係、斉次多項式に有効な adding variables method を利用して、バウンドを得た。Restricted Boltzmann Machine における研究では、先行研究 ``Cueto, Morton and B. Sturmfels, Geometry of the restricted Boltzmann machine'' および ``G. Montufar and N. Ay, Refinements of Universal Approximation Results for Deep Belief Networks and Restricted Boltzmann Machines'' との比較を行った。また ``Rusakov and Geiger, Asymptotic Model Selection for Naive Bayesian Networks'' の証明法を改良した。平成 24 年 6 月アメリカ、ペンシルベニア州立大学で開かれた国際会議 ``Algebraic Statistics in the Alleghenies(代数統計学)'' にて、題目 ``Learning coefficient and singular fluctuation in statistical learning theory'' について招待講演をした。

25 年度：Log canonical threshold の値を得るのに有用な 2 つの定理を得た。一つ目の定理「Method for finding a deepest singular point」は、log canonical threshold を得る最良の特異点を求めるものである。この定理は、ブローアッププロセス数の減少に貢献した 2 つ目の定理「Method to add variables」は、変数の追加により、ブローアッププロセス内で行われる変数変換を簡略化するのに有効であった。これら 2 つの定理は、斉次方程式からなるイデアルの特異点集合に効果がある。論文では、斉次イデアルから生成される Vandermonde matrix 型特異点に適用した。これらの結果は、Consideration on Singularities in Learning Theory and the Learning Coefficient, Entropy において発表した。Restricted Boltzmann Machine の学習係数の結果は、Learning coefficient in Bayesian estimation of restricted Boltzmann machine, Journal of Algebraic Statistics に掲載された。

26 年度：超関数を用いて特異学習モデルの尤

度関数である確率的複雑さの挙動を分析する定理において、別証明を与えた。学習理論において、超関数のメリン変換はゼータ関数であり、ゼータ関数の極はベイズ推定における学習モデルの尤度関数である確率的複雑さの挙動を与える。挙動の最大の次数を求める定理において、ヘルダーの不等式を用いて評価を行った。この結果は、博士前期課程高橋との共同研究であり、「学習理論における特異性の超関数を使った解析について」と題して、日本大学理工学部学術講演会にて発表した。また、学習誤差と関連する経験カルバック距離(経験過程)の挙動を考察した。特異点解消および中心極限定理を適用し、法則収束関数を求めることにより、経験過程の共通の標準形式を得る。これらをまとめたものを博士前期課程高村とともに、「学習理論における特異性と特異点解消について」と題し、日本大学理工学部学術講演会にて発表した。

27 年度：ARD 法を適用した場合の学習係数(log canonical threshold)について考察した。特に、脳活動の計測における MEG (Magnet oencephalography) 線形モデルに ARD 法を適用した場合の学習係数を求めた。脳活動の計測においては fMRI, EEG, MEG といった様々な計測方法が行われているが、MEG とは脳磁図・脳磁計と呼ばれ、脳内神経活動により発生する磁場を観測し神経活動を観測する装置である。この結果を岡田憲相とともに「学習理論と学習係数」と題し、RIMS Symposium(再生核の応用についての総合的な研究)で発表した。また、Vandermonde matrix 型特異点の log canonical threshold について考察を行なった。特異点を決定しているイデアルについて高次の項の影響を精密に調べ、中間層に対応するパラメータ数を低次元に固定した条件のもとで、結果を得ることが出来た。これは「Learning coefficients and reproducing true probability functions in learning systems」と題し、The tenth ISAAC Congress にて発表した。

28 年度：Vandermonde matrix 型特異点集合の log canonical threshold についてニューラルネットワークなどの中間層に対応するパラメータ数を低次元に固定した条件のもとで厳密な値を求め、その得られた結果を「Learning coefficients and reproducing true probability functions in learning systems」と題して、New Trends in Analysis and Interdisciplinary Applications, Trends in Mathematics(Springer)に掲載することになった。現在、イデアルを用いた特異点解消を用いて、効率よく値を得ることをもっと一般化された結果を導出することを目的として研究を継続している。また、修士 2 年生の岡田憲相とともに、学習により得られた確率密度関数と真の確率密度関数との

ずれを考察することを目的とした、ベイズ推定及びギブズ推定における汎化誤差・学習誤差間に成立する関係式について、すでに得られている結果の紹介、および、それらの別証明を与えたものを、日本大学工学部学術講演会にて「ベイズ推定とギブズ推定」と題して発表した。

#### 5. 主な発表論文等

(研究代表者, 研究分担者及び連携研究者には下線)

[雑誌論文](計 8 件)

青柳美輝, Learning coefficients and reproducing true probability functions in learning systems, New Trends in Analysis and Interdisciplinary Applications, Trends in Mathematics, 査読有, 2017, to appear, DOI:10.1007/978-3-319-48812-7\_44

青柳美輝, 岡田憲相, ベイズ推定とギブズ推定, 日本大学工学部学術講演会予稿集, 査読無, 2016, pp. 1200 - 1201.

青柳美輝, 岡田憲相, 学習理論と学習係数, 数理解析研究所講究録, 再生核の応用についての総合的な研究, 査読無, 2015, No.1980, pp. 220 - 233.

青柳美輝, 高村京輔, 学習理論における特異性と特異点解消について, 日本大学工学部学術講演会予稿集, 査読無, 2014, pp. 1307 - 1308.

青柳美輝, 高橋健彰, 学習理論における特異性の超関数を使った解析について, 日本大学工学部学術講演会予稿集, 査読無, 2014, pp. 1311 - 1312.

青柳美輝, Consideration on Singularities in Learning Theory and the Learning Coefficient, Entropy, 査読有, 2013, vol. 15, No. 9, 3714 - 3733, DOI:10.3390/e15093714.

青柳美輝, Learning coefficient in Bayesian estimation of restricted Boltzmann machine, Journal of Algebraic Statistics, 査読有, 2013, vol. 4, No. 1, 30 - 57

青柳美輝, Kenji Nagata, Learning coefficient of generalization

error  
in Bayesian estimation and Vandermonde matrix type singularity, Neural Computation, 査読有, 2012, vol. 24, No. 6, 1569 -1610.

[学会発表](計 5 件)

青柳美輝, 岡田憲相, 学習理論と学習係数, RIMS Symposium, 再生核の応用についての総合的な研究, 京都大学(京都府京都市), 平成 27 年 10 月 9 日

青柳美輝, Learning coefficients and reproducing true probability functions in learning systems, The tenth ISAAC Congress, Macau, China, 平成 27 年 8 月 5 日

青柳美輝, Learning coefficient and singular fluctuation in statistical learning theory, Algebraic Statistics in the Alleghenies, The Pennsylvania State University, Pennsylvania, USA, 平成 24 年 6 月 13 日

青柳美輝, Consideration on singularities in learning theory and real log canonical threshold, Singular learning theory: connecting algebraic geometry and model selection in statistics, AIM Research Conference Center (ARCC), Palo Alto, USA, 平成 23 年 12 月 15 日

青柳美輝, 学習理論と Vandermonde Matrix 型特異性, 第 55 回日本大学工学部学術講演会, 日本大学工学部(東京都千代田区), 平成 23 年 11 月 26 日

#### 6. 研究組織

(1) 研究代表者

青柳 美輝 (AOYAGI, Miki)  
日本大学・工学部・准教授  
研究者番号: 90338434