

科学研究費助成事業 研究成果報告書

平成 27 年 5 月 27 日現在

機関番号：12608

研究種目：基盤研究(C)

研究期間：2011～2014

課題番号：23500172

研究課題名(和文) 構造学習理論と双有理幾何学

研究課題名(英文) Structure Learning Theory and Birational Geometry

研究代表者

渡邊 澄夫 (WATANABE, SUMIO)

東京工業大学・総合理工学研究科(研究院)・教授

研究者番号：80273118

交付決定額(研究期間全体)：(直接経費) 2,600,000円

研究成果の概要(和文)：統計的学習においては、与えられたサンプルに対して確率モデルと事前分布の適切さを評価する規準としてベイズ自由エネルギーが知られているが、事後分布が正規分布で近似できないときにはその値を計算する方法は作られていなかった。本研究では、広く使えるベイズ情報量規準(WBIC)という概念を創出し、双有理幾何学に基づく数学的方法を応用することによりWBICがベイズ自由エネルギーと漸近的に同じ挙動を持つ確率変数であることを証明した。得られた定理を用いて、サンプルから階層構造を抽出する確率モデルにおいても統計的に適切なモデルを選ぶことができるようになった。

研究成果の概要(英文)：In statistical machine learning, it is well known that the appropriate model and prior for a given set of training samples is chosen by minimization of the Bayesian free energy. However, there has been no method to estimate the Bayesian free energy if the posterior distribution can not be approximated by any normal distribution. In this research, we created a new concept, a widely applicable Bayesian information criterion (WBIC), and proved that WBIC has the same asymptotic behavior as the Bayesian free energy, based on the birational geometry. The obtained theorem enables us to choose the optimal model for a given set of training samples, even if the model has hierarchical structures.

研究分野：数理情報学

キーワード：双有理幾何学 構造学習理論 ベイズ自由エネルギー WBIC 双有理不変量

1. 研究開始当初の背景

統計的学習には階層構造を持つモデルやモジュールを統合するタイプの情報処理が用いられることが多い。例えば、神経回路網、混合正規分布、ボルツマンマシン、隠れマルコフモデル、ベイズネットワークなどがその例である。これらの確率モデルにおいては、尤度関数を正規分布で近似することができないため、従来の統計理論によってその挙動を解析することができなかった。例えば、サンプルが与えられたときの確率モデルと事前分布の適切さを表す量であるベイズ自由エネルギーの値を知るための方法を作ることができなかった。

2. 研究の目的

事後分布が正規分布で近似できない場合には、確率モデルの持つ特異点の性質を数学的に扱う必要があり双有理幾何学が有用であることが知られている。本研究では、双有理幾何学の定理を応用して、事後分布が正規分布で近似できていなくても、真の分布が確率モデルで実現できていなくても、構造を持つ確率モデルの挙動を解析するための方法を確立する。

3. 研究の方法

従来の研究によって、ベイズ自由エネルギーの漸近挙動は双有理不変量である対数閾値とその多重度を用いて表す公式が導出されていたが、対数閾値とその多重度は真の分布に依存する値であり、統計的学習においては真の分布が不明であるため、従来の公式をそのまま実問題に適用することができなかった。本研究では真の分布が不明である場合においても対数閾値を推測する方法を構成し、ベイズ自由エネルギーを推測する公式を導出した。

4. 研究成果

事後分布が正規分布で近似できていなくても、真の分布が確率モデルで実現できていなくても、与えられたサンプルから自由エネルギーを推測することができる新しい情報量規準 WBIC を導出し、WBIC がベイズ自由エネルギーと等しい漸近挙動を持つことを証明した。以下にその概略を述べる。

情報源を表すある確率分布に独立に従う確率変数 X_1, X_2, \dots, X_n とパラメータ w を持つ確率モデル $p(x|w)$ と事前分布 (w) が与えられたとき、対数損失関数を

$$L(w) = -(1/n) \log p(X_i|w)$$

と定義し、ベイズ自由エネルギー F を

$$F = -\log \int \exp(-nL(w)) (w) dw$$

によって定義する。この値は、 n 個の確率変数が与えられたときの確率モデルと事前分布の対数尤度の符号反転である。このことから、 F が小さいほど適切なモデリングがなされていると考えることができるので統計的学習における重要な指標である。事後分布が正規分布で近似できる場合には、パラメータ w についての積分を正規分布の積分でおきかえることで F の近似値を求めることができるが、階層構造を持つ確率モデルにおいては関数 $L(w)$ の値が小さくなる領域が複雑な形をしているために正規分布で置き換える方法は適用できないことが知られていた。本研究では、逆温度が $1/\log n$ の事後分布による平均操作を $E[\cdot | 1/\log n]$ で表すとき、情報量規準 WBIC を

$$WBIC = -E[n L(w) | 1/\log n]$$

と新しく定義した。その上で、次の定理が成り立つことを証明した。

定理：真の分布が確率モデルで実現可能であってもなくても、事後分布が正規分布で近似できるときでもできないときでも、次の式が成り立つ。

$$F = WBIC + O_p((\log n)^{1/2})$$

ここで $O_p(f(n))$ は n が大きくなるとき $f(n)$ のオーダー以下の確率変数を表す記号である。

WBIC は逆温度が $1/\log n$ の事後分布をひとつ作ることで算出することができるので、上記の公式から計算が困難であったベイズ自由エネルギーの値を推測することが可能になった。

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4件)

中村文士, 渡邊澄夫, 一般ディリクレ分布を用いた混合正規分布の変分自由エネルギーの漸近挙動について, 電子情報通信学会論文誌 D, 査読有, Vol. J97-D, No.5, pp.1001-1013, 2014.

Sumio Watanabe, A widely applicable Bayesian information criterion, Journal of Machine

Learning Research, 査読有, Vol.14, pp.867-897, 2013.

Koshi Yamada, Sumio Watanabe, Statistical Learning Theory of Quasi-Regular Cases, IEICE Transactions, 査読有, Vol.E95-A, No.12, pp.2479-2487, 2012.

Daisuke Kaji, Sumio Watanabe, Two Design Methods of Hyperparameters in Variational Bayes Learning for Bernoulli Mixtures, Neurocomputing, 査読有, Vol.74, No.11, pp.2002-2007, 2011.

[学会発表](計 11件)

渡邊澄夫, 尤度関数がガウス近似できないときの統計的学習の評価指標について, 2014年7月24日-26日, 電子情報通信学会音声研究会, 花巻温泉ホテル, 岩手県花巻市.

Sumio Watanabe, Birational Invariants in Marginal Likelihood Computation, 2014 NIMS Thematic Program on Algebraic Statistics, 2014年7月14日~17日, National Institute of Mathematical Sciences, Daejeon, South Korea.

Sumio Watanabe, Discovery Phenomenon and Information Criteria, The Seventh Workshop on Information Theoretic Methods in Science and Engineering, 2014年7月5日~7日, Sheraton Waikiki Hotel, Honolulu, USA.

渡邊澄夫, 学習において一般漸近理論と正則漸近理論が成立するために十分なサンプル数の相違について, 電子情報通信学会NC研究会, 2013年12月21日, 岐阜大学, 岐阜県岐阜市.

玉井雄介, 渡邊澄夫, 最急降下法とMCMC法による情報量規準WBICに基づく階層型ニューラルネットワークのモデル選択, 電子情報通信学会IBISML研究会, 2013年11月12日, 東京工業大学, 東京都目黒区.

Sumio Watanabe, WAIC and WBIC are Information Criteria for Singular Statistical Model Evaluation, The Sixth Workshop on Information Theoretic Methods in Science and Engineering, Proceedings of WITMSE 2013, pp.90-94, 2013年8月26日~29日, The University of Tokyo, Japan.

Sumio Watanabe, Model Selection for Non-Regular Statistical Models, 2013 Joint Statistical Meetings, 2013年8月3日~8日, Palais des Congres de Montreal, Canada.

Sumio Watanabe, Resolution of Singularities and Statistical Model Evaluation, SIAM Conference on Applied Algebraic Geometry, 2013年8月1日~4日, Colorado State University, Fort Collins, Colorado, USA.

Koshi Yamada, Sumio Watanabe, Information Criterion for

Variational Bayes Learning in Regular and Singular Cases, Proc. of The 6th International Conference on Soft Computing and Intelligent Systems, F2-55-3, 2012年9月20日～24日、神戸国際会議場、兵庫県神戸市.

Sumio Watanabe. Algebraic geometry and model selection, Workshop of Singular Learning Theory : connecting algebraic geometry and statistical model selection, 2011年12月13日. American Institute of Mathematics, San Jose, CA, USA.

渡邊澄夫, 実現可能でなく正則でない場合の学習理論, 電子情報通信学会NC研究会, 2011年7月25日, 神戸大学, 兵庫県神戸市.

〔図書〕(計2件)

渡邊澄夫, 永尾太郎, 樺島祥介, 田中利幸, 中島伸一、ランダム行列の数理と科学, 森北出版, 2014.

渡邊澄夫, ベイズ統計の理論と方法, コロナ社, 2012.

〔産業財産権〕

出願状況(計 0 件)

取得状況(計 0 件)

〔その他〕

ホームページ等

<http://watanabe-www.math.dis.titech.ac.jp/users/swatanab/index-j.html>

6. 研究組織

(1) 研究代表者

渡邊 澄夫 (WATANABE SUMIO) 東京工業大学・大学院総合理工学研究科・教授

研究者番号: 80273118