

機関番号：17102

研究種目：基盤研究 (B)

研究期間：2007～2010

課題番号：19300051

研究課題名 (和文) 情報量概念を基盤とした学習理論の展開

研究課題名 (英文) Development of Learning Theory based on Information Measure

研究代表者

竹内 純一 (TAKEUCHI JUNICHI)

九州大学・大学院システム情報科学研究所・教授

研究者番号：80432871

研究成果の概要 (和文)：機械学習，情報理論，およびそれらの応用に関する諸課題について，記述長最小原理に基づく統一的視点のもとに研究を行った．特に，Markov モデルの幾何学的構造と確率的複雑度の関係，通信路容量と確率的複雑度の関係について考察し新たな知見を得た．また，アンサンブル学習等に関して考察し，効率的アルゴリズムや推定法を提案した．さらに，これら基礎的知見に基づき，ネットワークセキュリティにおけるインシデント検知，ポートフォリオ(分散投資戦略)，超解像などについて，新たな学習手法を提案し，その有効性を示した．

研究成果の概要 (英文)：We studied machine learning, information theory, and other related topics from a unified viewpoint of minimum description length principle (MDL principle). In particular, we obtained the new sight on the relation between geometrical structure of tree models and stochastic complexity (SC) and that between communication channel capacity and SC. We also studied ensemble learning and kernel method, and obtained efficient learning method for them. Further, on the basis of the fundamental knowledge obtained in this research, we proposed new learning based methods for incident detection in network security, universal portfolio, super resolution etc., and showed their efficiency.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2007年度	3,700,000	1,110,000	4,810,000
2008年度	3,500,000	1,050,000	4,550,000
2009年度	3,400,000	1,020,000	4,420,000
2010年度	3,700,000	1,110,000	4,810,000
総計	14,300,000	4,290,000	18,590,000

研究分野：学習理論

科研費の分科・細目：情報学・知能情報学

キーワード：学習理論，情報量，情報幾何，機械学習

1. 研究開始当初の背景

多量のデータから知識を抽出するための学習技術が重要性を増すなか，情報理論に端を発し，情報量概念を基盤とした学習理論の流れがある．すなわち，1978年に J. Rissanen が提唱した記述長最小原理 (Minimum Description Length Principle= MDL 原理) を核として広がる研究の流れである．MDL 原理とは，与えられたデータの中に規則性を見出すこと=学習は，その規則性(情報源のモ

デル)を利用してデータをコンパクトにまとめること，すなわちデータを圧縮することに通じるという思想である．これは「与えられたデータに対して同じ記述力を持つ複数の法則があれば単純なものほど予言力が高い」という Occam の剃刀の具体化である．具体的には，全記述長=モデル記述長+データ記述長が最小になるモデルを選ぶ方法を MDL 規準と呼ぶ．さらに，Rissanen はこの概念を推し進め確率的コンプレキシティ (Stochastic

Complexity = SC) の概念を確立した。SC は「何らかの方法でモデルを利用してデータを符号化する際の最小符号長」として定義された。

SC は、統計的モデル選択、逐次予測などで重要な役割を果たすことが知られており、世界中で盛んに研究されている。特に日本においては、MDL 規準に先駆けて、AIC (赤池の情報量規準) が 1974 年に提唱されたこともあり、この分野に関する関心が高い。

これらに関連した領域について、ニューラルネットや人工知能を含んだもう少し広い対象を研究する情報論的学習理論 (Information-Based Induction Sciences = IBIS) ワークショップが 1998 年から開催されている。IBIS には、情報理論、ニューロサイエンス、統計、統計物理、パターン認識、音声認識、自然言語処理、人工知能などをバックグラウンドとして多くの研究者が集まり、学際的な場を形成している。具体的テーマとしては、SC の評価、特異モデルの解析、確率伝播法、サポートベクターマシン (SVM)、アンサンブル学習などがある。

2. 研究の目的

本研究では、上記の課題について情報量概念を基盤としたパースペクティブを与えるとともに、有効な学習手法の構築することを目的とする。すなわち、上記諸課題を MDL 原理の観点から整理し、情報論的学習理論の特色をより明確にする成果を得る。同時に、関連する周辺分野との相互交流を行う。

3. 研究の方法

SC の性質に関する基礎研究と、それによって得られた知見に基づく応用研究を平行して進めた。ここでの応用研究は、関連する基礎分野と、学習技術の産業応用の両者を指している。具体的には以下の諸課題の研究を行った。

(1) Markov モデルの情報幾何

Markov 情報源の表現系である木情報源モデルの幾何学構造を明らかにした。

(2) SC とデータ圧縮

SC を達成する minimax 学習アルゴリズムや、種々のデータ圧縮方式の冗長度について考察した。

(3) アンサンブル学習等

アンサンブル学習やカーネル法等について、情報量概念の立場から考察した。

(4) SC と通信路容量の関係

期待値版 SC は通信路容量の関係を情報幾何の観点などから解析した。

(5) 機械学習の応用

ネットワークセキュリティ、ITS、カオス、超解像、ポートフォリオ (分散投資) その他について、機械学習の観点から取り組んだ。

4. 研究成果

(1) Markov モデルの情報幾何

木情報源モデルの幾何学的構造を明らかにした。Markov 連鎖モデルの状態集合について木構造を仮定することにより、その部分空間を定義できる。このように定義した情報源モデルは木情報源モデルと呼ばれ、Markov 連鎖モデルの一般化とみなせる。Markov 連鎖モデルのもう一つの一般化である FSM (Finite State Machine) モデルは漸近的に指数型分布族になることが知られている。一般の木情報源モデルは FSM ではないが、木がある条件を満たせば FSM となる。FSM となる木情報源モデルを FSMX モデルと呼ぶ。FSMX でない木情報源モデルが指数型に漸近するかどうかは不明であったが、本研究において指数型に漸近することを証明した。すなわち、木情報源モデルについては、指数型に漸近することと FSMX であることが同値であることを示した。

一方、モデルが i.i.d. の指数型分布族のとき、かつそのときに限り **Jeffreys 符号** (Jeffreys 事前分布を用いた **Bayes 符号**) が漸近的に SC を達成することが分かっていた。本研究では、このことが **Markov 情報源** の場合も同様に成り立つことを証明した。すなわち、木情報源モデルについても、それが指数型であるときに限って **Jeffreys 符号** によって漸近的に SC が達成できることを証明した。先に述べたこととあわせると、木情報源モデルはそれが **FSMX** であるときに限って **Jeffreys 符号** が漸近的に SC を達成する。これらの命題は図

	状態遷移関数	指数曲率	Jeffreys 符号
FSMX	存在 	ゼロ (自己平行) 	SC 達成
non FSMX	存在しない 	非ゼロ (曲面) 	SC 達成不可

図 1 木情報源モデルの分類

1 にまとめられる。このことは、木情報源モデルについては、グラフ的性質、幾何学的性質、情報量的性質それぞれが互いに関連していることを示している。

これらに関連して、さらに Markov 連鎖モデルの Fisher 情報量の性質を調べた。SC の定数項評価には、Fisher 情報行列の行列式を求めることが必要である。Markov 連鎖モデルの場合、Markov カーネルパラメータに関する Fisher 情報行列の行列式の計算は容易であるが、より重要な期待値パラメータについては、行列式の表式が知られていなかった。この問題に対して、Markov 連鎖モデルを包含する拡大モデルの Fisher 情報行列を利用することで、この形を求めることに成功した。この過程で、Markov 連鎖モデルの拡大モデルと

その Fisher 情報行列を定義したが、Fisher 情報行列が退化しており、奇妙な構造をもつことを明らかにした。

(2) SC とデータ圧縮

Context Tree Weighting (CTW) 法および Lemple-Ziv (LZ) アルゴリズムについて解析した。特に後者について、Markov 情報源向けに修正された LZ78 アルゴリズムの冗長度を評価した。また、特異事前分布を用いたユニバーサル符号の符号長を解析し、従来得られていた符号長の上界評価式を改善する事前分布を提案し理論的な保証を与えた。

(3) アンサンブル学習等

SVM の効率的アルゴリズムとして知られる分割法の解析を行った。分割法を基に凸 2 次計画問題の新たな解法アルゴリズムを提案し、数値実験によってその有効性を確認した。これは、SVM アルゴリズムの改善のみならず、一般の最適化手法に資するものである。

Boosting は弱学習器をモデルの外にシフトして改良する手法とみなせることに着目し、漸近的な解析を行った。モデルが真の分布を含むケースに、実際のシフトが理想的なシフトと正反対になることを示した。また、これを利用して、シフトを逆向きにすることで理想的シフトの近似とする手法を提案し、最尤推定値を改善できる場合があることを実験的に示した。

(4) SC と通信路容量の関係

パラメトリックな確率モデルは、パラメータを入力とし確率変数を出力とする通信路とみなせる。この通信路の容量は、確率モデルに関する期待値版 SC と密接に関連する。本研究ではこの関連を主に情報幾何の観点から考察した。特に、この関係は KL ダイバージェンスを Bregman ダイバージェンスに置き換えても同様に成り立つことを指摘した。また、離散通信路の容量を計算する室賀アルゴリズムを拡張し、期待値版 SC の上界を示した。

(5) 機械学習の応用

ネットワークセキュリティにおけるインシデント検知、ITS (高度交通システム) における旅行時間予測、超解像、一般物体認識、地震波解析、統計的手法による共通鍵暗号の解読、ユニバーサルポートフォリオなどに取り組んだ。

インシデント検知については、グラフィカルモデルのスパース学習を応用した手法により、現在最も大きな問題であるボットネットを検知できることを実験的に示した。

ポートフォリオに関しては、Cover のユニバーサルポートフォリオとユニバーサルデータ圧縮との関連を SC の観点から調べ、これまで知られていなかった幾何学的性質などを明らかにした。

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 9 件)

- ① Takeuchi, “Stochastic complexity, channel capacity, and universal portfolio,” Vol. 2, No. B, Journal of Math-for-Industry, pp. 213-225, 2010
- ② Takahashi, Ishitobi, Nishi, “Sufficient Conditions for One-Dimensional Cellular Neural Networks to Perform Connected Component Detection,” Nonlinear Analysis: Real World Applications, Vol. 11, pp. 4202-4213, 2010
- ③ Hafez, Khan, Kohda, “Earthquake onset detection using spectro-ratio on multi-threshold time-frequency sub-band,” Digital Signal Processing: A Review Journal, Vol. 19, pp. 118-126, 2009
- ④ Takahashi, Guo, Nishi, “Global Convergence of SMO Algorithm for Support Vector Regression,” IEEE Trans. Neural Networks, Vol. 19, pp. 971-982, 2008
- ⑤ Kawakita, Eguchi, “Boosting method for local learning in statistical pattern recognition,” Neural Computation, Vol. 20, pp. 2792-2838, 2008
- ⑥ Takeuchi, Kawabata, Barron, “Properties of Jeffreys mixture for Markov sources,” IEEE trans. Inform. Theory, (2007 年受理)

[学会発表] (計 47 件)

- ① 清田、奥村、川喜田、高橋、竹内, “ベイズ超解像における SIFT を用いたパラメータ推定手法の提案,” 電子情報通信学会技術研究報告 IE2010-168, 2011, 長崎県
- ② 濱崎、川喜田、竹内、井上、衛藤、中尾, “スパース構造学習を用いたボットネット検知法の提案とダークネットデータへの適用,” 2011 年暗号と情報セキュリティシンポジウム, 福岡県北九州市
- ③ Kawakita, Oie, Takeuchi, “A note on model selection for small sample regression,” International Symposium on Information Theory and its Applications, 2010, Taichung, Taiwan
- ④ Kawabata, “A Revisit to The Muroga Method of Computing Channel Capacity,” International Symposium on Information Theory and its

- Applications, 2010, Taichung, Taiwan
- ⑤ Kawakita, Takeuchi, “Semi-supervised learning in view of estimating functions,” Information Geometry And Its Applications III, 2010, Leipzig, Germany
- ⑥ 竹内, 池田, “通信路容量に関する情報幾何学的考察,” 第 33 回情報理論とその応用シンポジウム, 2010, 長野県
- ⑦ 鶴崎, 竹内, “ポートフォリオの Fisher 情報量について,” 第 33 回情報理論とその応用シンポジウム, 2010, 長野県
- ⑧ 前島, 實松, 香田, “揺らぎのあるニューラルネットワークにおける連想記憶,” 第 33 回情報理論とその応用シンポジウム, 2010, 長野県
- ⑨ 橋本, 川端, “特異事前分布をもつユニバーサル情報源符号の冗長度解析,” 電子情報通信学会情報理論研究会, 2010, 長野県長野市
- ⑩ 香田, 深江, “ β 展開を用いた AD/DA 変換の量子化誤差,” 電子情報通信学会情報理論研究会, 2010, 長野県長野市
- ⑪ Kawakita, Takeuchi, “A new method of improving predictive distribution without Bayes estimation,” 第 32 回情報理論とその応用シンポジウム, 2009, 山口県山口市
- ⑫ Takeuchi, “Fisher Information determinant and stochastic complexity for Markov Models,” Proc. of 2009 IEEE Intl Symp on Inform Theory, 2009 年 6 月, 韓国ソウル市
- ⑬ Inoue, Yoshioka, Eto, Yamagata, Nishino, Takeuchi, Ohkouchi, Nakao, “An Incident Analysis System nicter and Its Analysis engines Based on Data Mining Techniques,” the 15th Int. Conf. on Neural Information Processing, 2008, Auckland, New Zealand
- ⑭ 柿原, 竹内, 藤田, 姚, 岡, “ウェーブレット変換を用いた旅行時間時系列の予測,” 第 11 回情報論的学習理論ワークショップ, 2008, 宮城県仙台市
- ⑮ Kawabata, “Enumerative Implementation of Lempel-Ziv 77 Algorithm,” Proc. of 2008 IEEE International Symposium on Information Theory, 2008, Ontario, Canada
- ⑯ Guo, Takahashi, “Global Convergence Analysis of Decomposition Methods for Support Vector Regression” 2008 Int. Sympo. on Neural Networks, 2008, Beijing, China
- ⑰ Takeuchi, Kawabata, “Exponential

curvature of Markov models,” Proc. of 2007 IEEE Intl Symp. Inform. Theory, 2007 年 6 月, Nice, France.

〔図書〕(計 1 件)

- ① 小西貞則, 竹内純一(著, 分担執筆), 若山正人(編), 統計的モデリング/情報理論と学習理論, 講談社, 2008 年, pp. 63-116.

〔産業財産権〕

○取得状況(計 5 件)

名称: 変化点検出装置, 変化点検出方法および変化点検出用プログラム

発明者: 竹内純一, 山西健司

権利者: 日本電気株式会社

種類: 特許

番号: 特許公開 2 0 0 5 - 0 0 4 5 6 8

取得年月日: 2009 年 01 月 27 日

国内外の別: 国内

〔その他〕

ウェブサイト

<http://www-kairo.csce.kyushu-u.ac.jp/~tak/homepage.ja.html>

6. 研究組織

(1) 研究代表者

竹内 純一 (TAKEUTHI JUNICHI)

九州大学・大学院システム情報科学研究
院・教授

研究者番号: 8 0 4 3 2 8 7 1

(2) 研究分担者

高橋 規一 (TAKAHASHI NORIKAZU)

九州大学・大学院システム情報科学研究
院・准教授

研究者番号: 6 0 2 8 4 5 5 1

實松 豊 (JITSUMATSU YUTAKA)

九州大学・大学院システム情報科学研究
院・准教授

研究者番号: 6 0 3 3 6 0 6 3

川端 勉

電気通信大学・電気通信学部・教授

研究者番号: 5 0 1 5 2 9 9 7

川喜田 雅則

九州大学・大学院システム情報科学研究
院・助教

研究者番号: 9 0 4 3 5 4 9 6

香田 徹

九州大学・大学院システム情報科学研究
院・特任教授

研究者番号: 2 0 0 3 8 1 0 2

(H22: 連携研究者)