

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成 24 年 4 月 18 日現在

機関番号：34416

研究種目：基盤研究(C)

研究期間：2009～2011

課題番号：21500228

研究課題名（和文）

多様なモデルのオンライン学習に関する網羅的・系統的研究

研究課題名（英文）

Exhaustive and systematic analyses on online learning of various models

研究代表者

三好 誠司 (MIYOSHI SEIJI)

関西大学・システム理工学部・教授

研究者番号：10270307

研究成果の概要（和文）：学習機械の汎化能力をオンライン学習の枠組みで統計力学的手法を用いて解析した。解析は、生徒が複数の動く教師から学ぶモデル、学習に用いる例題のうち少数の例題だけに正解が付与されているようなモデル、例題自身に相関があるモデルなど、統計的学習の基礎理論と応用の両面への波及効果が大きく、時間的あるいは空間的な観点で特徴を有するいくつかのモデルを対象とし、網羅的かつ系統的に行った。また、いずれの場合も計算機実験との比較により検証を行った。

研究成果の概要（英文）：We have analyzed the generalization ability of learning machines in the framework of online learning using statistical mechanical method. The exhaustive and systematic analyses have treated some models which are interesting from the viewpoint of temporal and spacious aspects and have effects on both the fundamental theory and applications.

交付決定額

（金額単位：円）

| | 直接経費 | 間接経費 | 合計 |
|--------|-----------|---------|-----------|
| 2009年度 | 1,000,000 | 300,000 | 1,300,000 |
| 2010年度 | 1,000,000 | 300,000 | 1,300,000 |
| 2011年度 | 1,000,000 | 300,000 | 1,300,000 |
| 年度 | | | |
| 年度 | | | |
| 総計 | 3,000,000 | 900,000 | 3,900,000 |

研究分野：総合領域

科研費の分科・細目：情報学・感性情報学・ソフトコンピューティング

キーワード：情報統計力学，統計的学習理論，統計力学，機械学習，オンライン学習，汎化能力，汎化誤差

1. 研究開始当初の背景

与えられた一定数の入力と正解のペア（例題）を用いて、生徒が繰り返し学習する枠組みはバッチ学習と呼ばれる。これに対して、一つの例題が与えられたら、生徒はそれに応じてすぐに自分自身のパラメータ（結合荷重など）を変化させ、その例題は忘れてしまって二度と使わないような学習の枠組みはオ

ンライン学習と呼ばれる。オンライン学習は過去に出てきた例題についても必ずしも確実には答えられないという欠点があるものの、学習時間や記憶容量の点でメリットが大きい。また、教師の構造が時間とともに変化する場合などのように時間とともに環境が変化しても追従できる。オンライン学習を対象とした統計力学的解析については 1990 年代以降に、たとえば D. Saad, M. Oppen, M.

Biehl, P. Riegler, J. K. Anlauf, J. Inoue, Y. Kabashima, H. Nishimori らによって行われてきた。

申請者はこれらの先行研究の結果をふまえて 2006-2008 年度の科学研究費補助金により、教師が動くモデル、時間方向アンサンブル学習、教師が複数存在するモデルなど時間的・空間的な観点で興味深いモデルを対象に統計力学的手法を用いた基本的な解析を行った。その結果、これらが比較的シンプルなモデルであるにも関わらず非自明で興味深い性質を有することが明らかになった。たとえば、教師が動くモデルにおいては、生徒からは動く教師の出力だけが見えており、真の教師の出力は直接には見えないにも関わらず、ある条件においては生徒が動く教師よりも賢くなりうる。また、教師が複数存在するモデルにおいて生徒が賢くなるためには必ずしも教師の数が多かったり、その多様性が豊かである方が良いわけではない。

このように 2006-2008 年度の科研費による研究は統計力学的手法を用いた解析の第一歩と呼ぶべきものであるにも関わらず、当初期待した以上の成果を上げることができた。

2. 研究の目的

教師が動くモデルや教師が複数存在するモデルを対象とした網羅的解析、さらにこれらを拡張したモデルを対象とする発展的解析を行う。具体的な内容は以下の通りである。

(1) 生徒が複数の動く教師から学ぶモデルの解析

教師が動くモデルの解析と教師が複数存在するモデルの解析をもとに、生徒が複数の動く教師から学ぶモデルの解析を行う。現実の人間社会の階層構造においても教師は複数存在しており、また、それぞれの教師は自身が学び続けている場合が多いので、この解析は現実の人間社会との類似性を持っている。

(2) 新しい学習則

2006-2008 年度の研究で対象とした学習則は線形モデルの場合は勾配法、非線形モデルの場合はヘブ学習、パーセプトロン学習、アダプトン学習であった。しかし、学習則はもちろんこれだけではない。たとえば最近注目されている学習則のひとつに同時摂動型学習がある。このように従来からよく解析されている学習則以外の学習則の解析に取り組む。

(3) 入力に関連を有する場合

統計力学的手法を用いる場合、解析の都合上、入力に独立に生成されることを仮定する機会が多い。この仮定により解析は容易になるが、実際の応用を考えた場合、入力に相関を有する機会が多い。手書き文字認識を例にとると、例題パターンを多く収集する必要があるが、このとき同一人物が書く文字パターンは似ている（相関を有する）はずである。このように入力に相関を有する場合について理論的に解析しておくことは有意義である。

3. 研究の方法

(1) モデルの設定

教師や生徒は線形パーセプトロンあるいは非線形パーセプトロン（単純パーセプトロン）であるとする。結合荷重の初期値や入力はガウス分布にしたがい独立に生成されるものとする。また、本来は確率的なダイナミクスを決定論的な連立微分方程式で記述するために、結合荷重や入力の次元が無量大の**熱力学的極限**を考えるものとする。さらに、以下で述べる自己平均性を仮定するために、教師や生徒の結合荷重のノルムが入力のノルムより十分大きい場合を考える。

(2) 巨視的変数のダイナミクスを記述する決定論的な連立微分方程式の導出

モデルのダイナミクスを少数の巨視的変数で記述する。具体的には、生徒の長さや生徒と教師の類似度が巨視的変数となる。これらのダイナミクスを記述する連立微分方程式を統計力学的手法に基づき導出する。この研究で扱うモデルでは、教師や生徒の結合荷重のノルムが入力のノルムより十分大きいので、本来は確率変数である入力の影響をその平均で置き換えることができる。この考え方を統計物理では**自己平均性**と呼ぶ。その結果、連立微分方程式は決定論的な形で導出されることになる。

(3) サンプル平均の計算

導出された決定論的連立微分方程式にはいくつかのサンプル平均、すなわち確率変数を入力に関して平均したものが含まれている。これを解析的に計算する。線形パーセプトロンの場合は比較的容易に解析的に求めることができる場合が多い。非線形パーセプトロンの場合にはこのサンプル平均を計算するために複雑なガウス積分を実行しなければならない場合も多い。その場合には数値積分やモンテカルロ法を用いる。

(4) 微分方程式を解析的に解く

解析的に求められたサンプル平均を連立微分方程式に代入し、これを解くことにより、巨視的変数のダイナミクスが理論的に求められる。線形パーセプトロンの場合は連立微分方程式が解析的に解ける可能性も十分にある。非線形パーセプトロンの場合には仮にサンプル平均が解析的に求まっても、それらを代入した連立微分方程式は解析的に解けない場合も多いが、その場合にはルンゲクッタ法等を用いて数値的に解く。

(5) 汎化誤差の計算

統計的学習理論の重要な目的は、未知の入力に対する誤差の期待値、すなわち汎化誤差を理論的に計算することである。解析的に、あるいは数値的に求められた巨視的変数を用いて汎化誤差を計算する。ここでも線形パーセプトロンの場合は解析的に計算できる可能性が高いのに対し、非線形パーセプトロンの場合にはガウス積分を数値的に実行したりモンテカルロ法を用いる必要が生じる可能性もある。

(6) 計算機実験の実行

以上の理論解析と並行して計算機実験を行う。その場合、結合荷重や入力の次元を無限大とすることはできないので、 $10^3 \sim 10^4$ 次元で実行する。また、汎化誤差の計算のために $10^6 \sim 10^8$ 個程度のランダム入力を生成し、それらに対する誤差の平均をとる。

(7) 検証と考察

理論計算結果と計算機実験の結果を比較検討することにより理論の正当性を検証する。また、計画通り巨視的変数や汎化誤差が解析的に得られたならば、汎化誤差が発散しない条件や時間が十分経過した後の定常状態、残留汎化誤差についても理論的な洞察が行える可能性が高いので、検討・考察を行う。

4. 研究成果

(1) 生徒が複数の動く教師から学ぶモデルの解析

多数の非定常な教師から生徒がオンライン学習を行う場合について、教師と生徒がいずれも雑音が重畳する線形パーセプトロンであるようなモデルを対象に統計力学的手法を用いて解析的に調べた。その結果、教師の学習係数、生徒の学習係数、教師の数によって生じる汎化誤差の違いが定量的に明らかになった。また、多数の教師自身が学習をしている場合としていない場合について、生徒の汎化誤差に違いが生じることが明らかになった。

(2) 新しい学習則

オンライン型の同時摂動型学習の汎化誤差の漸近特性はパーセプトロン学習と同じで例題数の $-1/3$ 乗であることが先行研究により明らかになっているが、我々はパーセプトロン学習とアダルトン学習の違いを深く考察し、これを同時摂動型学習に適用することにより適応型同時摂動学習を提案し、その汎化能力を統計力学的手法を用いて解析した。その結果、提案学習則の汎化誤差がアダルトン学習と同様に例題数の -1 乗に比例する漸近特性を有することを明らかにした。

また、半教師付学習の解析を行った。ヘブ学習とパーセプトロン学習の二つの学習則を対象に解析した結果、ヘブ学習の場合には、同時に用いる入力パターンの数 K が大きいほど学習初期での汎化誤差減少は速いが、学習終盤では逆に K が小さい方が汎化誤差が小さくなることがわかった。一方、パーセプトロン学習の場合には汎化誤差は減少しないことが明らかになった。

(3) 入力に関連を有する場合

関連を有する例題を用いるオンライン学習の解析を行い、汎化誤差のダイナミクスを理論的に求めた。その結果、パーセプトロン学習では同時に用いる例題を増やすと汎化誤差が減少するのに対し、ヘブ学習とアダルトン学習では下限があり、学習則により質的な違いが存在することを示唆する結果が得られた。

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12件)

- ① Hara, K. and Miyoshi, S., Statistical mechanics of on-line ensemble teacher learning through a novel perceptron learning rule, Journal of the Physical Society of Japan, 査読有, (accepted)
- ② Nakao, K., Narukawa, Y., and Miyoshi, S., Statistical mechanics of on-line learning using correlated examples, IEICE Transactions on Information and Systems, Vol.E94-D, No.10, pp.1941-1944, Oct. 2011. 査読有
- ③ Miyoshi, R., Maeda, Y., and Miyoshi, S., Statistical mechanics of adaptive weight perturbation learning, IEICE Transactions on Information and

- Systems, Vol.E94-D, No.10, pp.1937-1940, Oct. 2011. 査読有
- ④ Hasegawa,R., Okada,M., and Miyoshi,S., Image segmentation using region-based latent variables and belief propagation, Journal of the Physical Society of Japan, Vol.80, No.9, 093802 (4 pages), Sep. 2011. [This paper was selected as the "Papers of Editors' Choice."] 査読有
- ⑤ Miyoshi,S. and Kajikawa,Y., Statistical-mechanics approach to the Filtered-X LMS algorithm, Electronics Letters, Volume 47, Issue 17, pp.997-999, Aug. 2011. 査読有
- ⑥ 山田貴博, 三好誠司, 肥川宏臣, 前田裕, 同時摂動を用いた PSO の FPGA 実現, 電気学会論文誌 C Vol.131, No.3, pp.682-688, March 2011. 査読有
- ⑦ Miyoshi,S. and Okada,M., Image restoration and segmentation using region-based latent variables: Bayesian inference based on variational method, Journal of the Physical Society of Japan, Vol.80, No.1, 014802 (11 pages) Jan. 2011. 査読有
- ⑧ Tsuchie,K., Hanada,Y., and Miyoshi,S., Estimation of distribution algorithm incorporating switching, IEICE Transactions on Information and Systems, Vol.E93-D, No.11, pp.3108-3111, Nov. 2010. 査読有
- ⑨ Uezu,T., Abe,K., Miyoshi,S., and Okada,M., Statistical mechanical study of partial annealing of a neural network model, Journal of Physics A: Mathematical and Theoretical, Vol.43, 025004 (22 pages), Jan. 2010. 査読有
- ⑩ Hara,K., Nakayama,Y., Miyoshi,S., and Okada,M., Statistical mechanics of on-line mutual learning with many linear perceptrons, Journal of the Physical Society of Japan, Vol.78, No.11, 114001 (5 pages), Nov. 2009. 査読有
- ⑪ Miyoshi,S., Hikawa,H., and Maeda,Y., Statistical mechanical analysis of simultaneous perturbation learning, IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, Vol.E92-A, No.7, pp.1743-1746, July 2009. 査読有
- ⑫ Miyoshi,S. and Okada,M., Effect of slow switching of ensemble teachers in on-line learning, Journal of the Physical Society of Japan, Vol.78, No.5, 053001 (4 pages), May 2009. 査読有
- [学会発表] (計 47 件)
- ① 木下俊貴, 岡田真人, 三好誠司, 複層ベイズ超解像の画像への適用, 日本物理学会年次大会 (関西学院大学), 27pAG-4, 2012.3.27
- ② 長谷川亮太, 瀧山健, 岡田真人, 三好誠司, スイッチング状態空間モデルと変分ベイズ法による画像の修復と領域分割, 日本物理学会年次大会 (関西学院大学), 27pAG-3, 2012.3.27
- ③ 加藤弘之, 岡田真人, 三好誠司, CDMA の SCSNA, 日本物理学会年次大会 (関西学院大学), 27aAG-10, 2012.3.27
- ④ 三好誠司, 梶川嘉延, 適応フィルタの統計力学 II, 日本物理学会年次大会 (関西学院大学), 27aAG-9, 2012.3.27
- ⑤ 三好誠司, 梶川嘉延, 統計力学的手法による Filtered-X LMS アルゴリズムの解析, 信号処理シンポジウム(札幌コンベンションセンター), C9-1 (6 pages), 2011.11.18
- ⑥ 長谷川亮太, 瀧山健, 岡田真人, 三好誠司, スイッチング状態空間モデルと変分ベイズ法による画像の修復と分割, 信号処理シンポジウム(札幌コンベンションセンター), A1-4 (6 pages), 2011.11.16
- ⑦ 三好誠司, 梶川嘉延, 統計力学的手法による適応信号処理の解析, 第 14 回情報論的学習理論ワークショップ (IBIS2011, 奈良女子大学) ディスカッショントラック D-136, 2011.11.10
- ⑧ 萱野健太, 永田賢二, 岡田真人, 三好誠司, 変分ベイズ法と MCMC を用いた画像の修復と領域分割, 第 14 回情報論的学習理論ワークショップ (IBIS2011, 奈良女子大学) テクニカルトラック T-27, 2011.11.10, 電子情報通信学会技術研究報告, Vol.111, No.275, IBISML2011-68, pp.175-180

- ⑨ 三好亮介, 前田裕, 三好誠司, 適応型荷重摂動学習の統計力学, 情報論的学習理論ワークショップ (IBIS2010. 東大駒場), 電子情報通信学会技術研究報告, IBISML2010-92, pp. 245-250, 2010. 11. 5
- ⑩ 中尾健人, 鳴川雄太, 三好誠司, 入力が相関を有するオンライン学習の統計力学, 情報論的学習理論ワークショップ (IBIS2010. 東大駒場), 電子情報通信学会技術研究報告, IBISML2010-91, pp. 239-244, 2010. 11. 5

[その他]

<http://www2.itc.kansai-u.ac.jp/~miyoshi/cnt/gyoseki.html>

6. 研究組織

(1) 研究代表者

三好 誠司 (MIYOSHI SEIJI)
関西大学・システム理工学部・教授
研究者番号: 10270307

(2) 研究分担者

なし

(3) 連携研究者

なし