

平成21年6月18日現在

研究種目：若手研究 (B)  
 研究期間：2006～2008  
 課題番号：18700157  
 研究課題名 (和文) 一般化した確率過程に基づくオンライン学習の漸近的解析  
 研究課題名 (英文) Asymptotic Analysis of On-line Learning  
 Based on Generalized Stochastic Processes  
 研究代表者  
 岩田 一貴 (IWATA KAZUNORI)  
 広島市立大学・大学院情報科学研究科・助教  
 研究者番号：20405492

## 研究成果の概要：

強化学習における経験系列 (状態, 行動, 報酬の時系列) やクラスタ分析に使われるデータは, 一般に非定常・非エルゴードな確率過程に従う. 本研究では, そのような一般的な系列・データに対しても有効な漸近的性質を使って, (1)マルチエージェントシステムの強化学習における漸近的性質, (2)強化学習によって収益の最大化が可能な確率的決定過程のクラス, (3)階層クラスタ分析におけるクラスタ間の新しい非類似度についての研究を行った.

## 交付額

(金額単位：円)

	直接経費	間接経費	合計
2006年度	1,300,000	0	1,300,000
2007年度	1,000,000	0	1,000,000
2008年度	1,000,000	300,000	1,300,000
年度			
年度			
総計	3,300,000	300,000	3,600,000

研究分野：総合領域

科研費の分科・細目：情報学・知能情報学

キーワード：確率的決定過程, 強化学習, マルチエージェントシステム, 階層クラスタ分析

## 1. 研究開始当初の背景

オンライン学習は人工知能の分野を中心に活発に研究されているテーマで, 例えば確率降下法, 強化学習, 適応フィルタなど多くの学習アルゴリズムがこれに分類される. それらのアルゴリズムは, それぞれパターン認識, ロボット制御, 音声処理などに応用されている. 国内ではほとんどの研究がオンライン学習の応用に関するもので, 本研究のような理論的立場からの基礎的研究はあまり見られない. 一方で, 国外では基礎的な研究も活発に行われ, 主なものとしては確率近似法に基づくものが挙げられる. これは情報源を正規

分布などのパラメトリックな確率分布で定義し, 時間軸上で局所的に定常性およびエルゴード性を仮定した上で, 確率分布の推定パラメータの時系列の軌跡を常微分方程式で表す手法である. この微分方程式が漸近的に推定パラメータの期待値に収束する条件を調べることで収束解析を行うことができ, その結果はオンライン学習を応用する際の大きな指針となっている. これらの研究はどの推定方法を選ぶか, どの収束の強さ (概収束か確率収束など) を保証するかでその結果が異なってくる. 本研究は, オンライン学習において定常性およびエルゴード性を仮定せ

ず、情報スペクトルに基づいたオンライン学習の確率過程の解析を行うもので、確率近似法に基づく解析と比べると、より一般的な立場からの研究と言える。

## 2. 研究の目的

訓練データを観測しながら逐次的に学習を行なうオンライン学習は、前もって訓練データを用意する必要がなく、ある程度の環境の変化にも追従できるために、音声処理やロボット制御、Web データマイニングにおいて良く使われる。従来のオンライン学習では、訓練データを生成する確率分布によって定義される情報源が定常かつエルゴード性を満たす場合、もしくはそのように見なせるほど情報源がゆっくり変化する場合に、どのように最適な解を学習していくのかということに注目してきた。しかし、実際の応用においては情報源の定常性やエルゴード性が厳密に成り立つ場合は稀である。例えば、ロボット制御などでは自分自身の行動の結果次第で環境の構造が変化していくために定常であるとはいえない。また、様々なグループから集められた Web データの情報源は、複数の情報源から成る混合情報源となり、このような情報源は一般的にエルゴード性を満たさない。そこで、本研究ではオンライン学習に使う訓練データが、非定常や非エルゴードな情報源から生成される場合や、訓練データにノイズが含まれる場合におけるオンライン学習の確率過程の漸近的な振る舞い（最適解への収束）を解析するのが目的である。その結果は、オンライン学習を使う様々な応用に大きな指針をもたらすと期待できる。

## 3. 研究の方法

具体的なオンライン学習として強化学習を取り扱う。強化学習は、観測される経験系列（状態、行動、報酬の時系列）に自分自身の学習（行動選択戦略の改善）の結果が徐々に反映され、最終的にその経験系列自体を最適化するという点で、本研究課題が扱う非定常・非エルゴードの確率過程に基づくオンライン学習の良い例になっているからである。また、もう一つの理由は、強化学習は研究代表者が熟知しているオンライン学習の枠組みであるということが挙げられる。そこで、従来の強化学習の枠組みである定常エルゴードマルコフ決定過程における漸近的性質をまとめることから始める。その後、枠組みをより一般的な非定常・非エルゴードの確率的決定過程に一般化し、そこから導かれる漸近的性質を明らかにする。パターン認識分野における階層クラスタ分析に対しても本研究のアプローチは有効である。標本の母集団が混合分布モデルである場合、各標本はそれぞれの部分母集団の確率

分布に従うので、このような標本の生成過程もまた一般に非定常・非エルゴードとなる。そこで、Sanov の定理を混合分布モデルに適用することで、情報源符号化および統計的仮説検定という意味で理論的に妥当な部分母集団の間の非類似度を導く。また、人工データおよび実データを使った階層クラスタ分析を行い、導いた非類似度が優れた分類結果を与えることを示す。

## 4. 研究成果

研究成果は次の (1), (2), (3) に分けられる。  
(1) マルチエージェントシステムの強化学習における漸近的性質  
マルチエージェントシステムとは、複数のエージェントが、互いに協調することにより、ある与えられたタスクに関して最適な政策を学習するような枠組みである。その枠組みが従う確率過程としては、Temporal Difference 学習などの代表的な強化学習アルゴリズムが適用し易いなどの理由から、マルコフ決定過程がよく使われる。よって、本研究では、定常エルゴードマルコフ決定過程に従うマルチエージェントシステムにおいて成立する漸近的性質を示し、最適政策の学習によりシステム全体の収益が最大化される過程を確率的な観点から明らかにした。この成果は、学習理論分野の主要な国際誌である IEEE Transactions on Neural Networks に掲載された（雑誌論文 4 参照）。

(2) 強化学習によって収益の最大化が可能な確率的決定過程のクラス  
強化学習とはエージェントと環境の相互作用の枠組みのことで、通常はある確率的決定過程によって記述される。エージェントは、その確率的決定過程に従い、行動選択-実行-状態遷移を繰り返しながら、収益と呼ばれるスカラー量を最大化するような行動選択の方法を学習する。相互作用の枠組みが定常エルゴードマルコフ決定過程に従うならば、ある意味で最適な行動選択の方法を学習するためのアルゴリズム (Q-learning など) が良く知られている。そのため、強化学習をある問題に対して応用する際には、相互作用の枠組みが定常エルゴードマルコフ決定過程に従うと仮定してしまうことがほとんどである。しかしながら、実際には、その相互作用の枠組みは定常性・エルゴード性・マルコフ性を満たしていないことが多い。にもかかわらず、そのような問題に対して、そもそも強化学習を応用することの有効性やある最適な行動選択の方法を学習するアルゴリズムが存在するかどうかについてはよくわかっていない。  
本研究では、情報理論的な視点から、漸近等分割性と収益最大化の関係を明らかにする

ことによって、収益の最大化が可能な確率的決定過程のクラスを導出した。また、そのクラスは定常エルゴードマルコフ決定過程よりも広いことがわかった。そのクラスに含まれないような確率的決定過程では収益最大化が不可能であるため、そのクラスはある問題に対して強化学習を応用することが有効であるかどうかを判断するのに有益な情報を与える。この成果は、人工知能分野の主要な国際会議である IAT2008 で発表した（研究発表 7 参照）。

(3) 階層クラスタ分析におけるクラスタ間の新しい非類似度  
階層クラスタ分析は、統計データから母集団のある傾向を分析するためによく使われる。本研究では、階層クラスタ分析における標本のラベル（分類値）の階層性をボトムアップ的に推定することを目的とした、新しいクラスタ（標本の集合）間の非類似度を提案した。この非類似度は、標本の母集団が混合モデルである場合に、ある種の情報理論的な妥当性を持っている。  
実際のデータに対するクラスタ分析の成果として、窓ガラスなどのガラス成分データや音声を構成する音素データの階層構造を同定する際に、提案した非類似度が有効であることを計算機実験により示した。これらの成果は、パターン認識分野の主要な国際誌である IEEE Transactions on Pattern Analysis and Machine Intelligence（雑誌論文 2 参照）および国内誌の電子情報通信学会論文誌に掲載された（雑誌論文 1 参照）。

#### 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）

1. 斧城 悠大, 岩田 一貴, 末松 伸朗, 林 朗, 確率密度推定に基づく RDSP法を用いた音素データの階層クラスタ分析, 電子情報通信学会論文誌, v.J91-D, 2008, pp.2196-2200, 査読有.
2. Kazunori Iwata and Akira Hayashi, A Redundancy-Based Measure of Dissimilarity among Probability Distributions for Hierarchical Clustering Criteria, IEEE Transactions on Pattern Analysis and Machine Intelligence, v. 30, 2008, pp. 76-88, 査読有.
3. 大内 悠, 岩田 一貴, 林 朗, スケールの影響に対して頑健な曲線の最適整合, 情報科学技術レターズ, v. 6, 2007,

pp. 209-210, 査読有.

4. Kazunori Iwata, Kazushi Ikeda and Hideaki Sakai, A Statistical Property of Multiagent Learning Based on Markov Decision Process, IEEE Transactions on Neural Networks, v. 17, 2006, pp. 829-842, 査読有.

〔学会発表〕（計 26 件）

1. Kenji Ono, Kazunori Iwata and Akira Hayashi, An Action-Selection Strategy Insensitive to Parameter-Settings in Reinforcement Learning, 08/2009, The ICROS-SICE International Joint Conference 2009, Fukuoka, Japan. (accepted)
2. Kazunori Iwata, On a Class of Stochastic Decision Processes with an Asymptotic Property, 電子情報通信学会 2009 年総合大会, 2009 年 3 月 18 日, 愛媛大学.
3. 上青木 勝利, 岩田 一貴, 末松 伸朗, 林 朗, 曲線像から相似な部分を抽出するための高速なアルゴリズム, 電子情報通信学会 2009 年総合大会 情報・システムソサイエティ特別企画学生ポスターセッション, 2009 年 3 月 17 日, 愛媛大学.
4. 岡田 行史, 末松 伸朗, 林 朗, 岩田 一貴, Poly Tree 階層モデルの実験的研究, 電子情報通信学会 2009 年総合大会 情報・システムソサイエティ特別企画学生ポスターセッション, 2009 年 3 月 17 日, 愛媛大学.
5. 金子 悟士, 林 朗, 岩田 一貴, 末松 伸朗, 時系列データの低次元ユークリッド空間への埋め込み, 電子情報通信学会 2009 年総合大会 情報・システムソサイエティ特別企画学生ポスターセッション, 2009 年 3 月 17 日, 愛媛大学.
6. 藤田 睦未, 林 朗, 岩田 一貴, 末松 伸朗, 変分ベイズ法の環境音認識への応用, 電子情報通信学会 2009 年総合大会 情報・システムソサイエティ特別企画学生ポスターセッション, 2009 年 3 月 17 日, 愛媛大学.
7. Kazunori Iwata, An Information-Theoretic Class of Stochastic Decision Processes, The 2008 IEEE/WIC/ACM International Conference on Intelligent Agent

- Technology, 11/12/2008, Sydney, Australia.
8. Kazunori Iwata and Akira Hayashi, Sampling Curve Images to Find Similarities among Parts of Images, The 15th International Conference on Neural Information Processing, 27/11/2008, Auckland, New Zealand.
  9. 岩田 一貴, 林 朗, 曲線整合において相似な部分を見つけるためのサンプリング方法, 第 11 回情報論的学習理論ワークショップ, 2008 年 10 月 30 日, 仙台国際センター.
  10. 小野 兼嗣, 岩田 一貴, 林 朗, 強化学習におけるソフトマックス戦略のパラメータの新しい調整法, 第 11 回情報論的学習理論ワークショップ, 2008 年 10 月 30 日, 仙台国際センター.
  11. 小野 兼嗣, 岩田 一貴, 林 朗, 強化学習におけるパラメータ設定に頑健な行動選択戦略, 第 7 回情報科学技術フォーラム, 2008 年 9 月 3 日, 慶應義塾大学 湘南藤沢キャンパス.
  12. 鷲頭 祐樹, 末松 伸朗, 林 朗, 岩田 一貴, 混合ディリクレ過程モデルを利用したARMAモデルベース時系列クラスタリング, 第 7 回情報科学技術フォーラム, 2008 年 9 月 3 日, 慶應義塾大学 湘南藤沢キャンパス.
  13. 斧城 悠大, 岩田 一貴, 末松 伸朗, 林 朗, RDSP法を使った音素データの階層クラスタ分析, 電子情報通信学会NC研究会, 2008 年 3 月 14 日, 玉川大学.
  14. Kazushi Ikeda and Kazunori Iwata, Information Geometry and Information Theory in Machine Learning, The 14th International Conference on Neural Information Processing, 15/11/2007, Kitakyushu, Japan.
  15. Kazunori Iwata and Akira Hayashi, Identifying the Underlying Hierarchical Structure of Clusters in Cluster Analysis, The 17th International Conference on Artificial Neural Networks, 10/9/2007, Porto, Portugal.
  16. 大内 悠, 岩田 一貴, 林 朗, スケールの影響に対して頑健な曲線の最適整合, 第 6 回情報科学技術フォーラム, 2007 年 9 月 5 日, 中京大学.
  17. 斧城 悠大, 岩田 一貴, 末松 伸朗, 林 朗, 確率密度推定を用いたRDSP法によるクラスタの階層構造の調査, 第 6 回情報科学技術フォーラム, 2007 年 9 月 5 日, 中京大学.
  18. 成田 浩之, 林 朗, 末松 伸朗, 岩田 一貴, 半正定値計画法による時系列データの埋め込み, 第 6 回情報科学技術フォーラム, 2007 年 9 月 5 日, 中京大学.
  19. 杉浦 玄朗, 林 朗, 後藤 直人, 末松 伸朗, 岩田 一貴, 階層隠れCRFの提案, 第 6 回情報科学技術フォーラム, 2007 年 9 月 5 日, 中京大学.
  20. 岩田 一貴, 林 朗, Cluster Analysis for Identifying the Hierarchical Structure of Subpopulations in Mixture Models, 電子情報通信学会PRMU研究会, 2007 年 9 月 4 日, 名古屋大学.
  21. 岩田 一貴, 強化学習における大偏差理論の応用, 計測自動制御学会中部支部講演会「学習理論と制御」, 2006 年 12 月 26 日, 名古屋大学, 招待講演.
  22. 岩田 一貴, 池田 和司, 酒井 英昭, 強化学習における漸近的性質と確率的な行動選択戦略との関係, 第 49 回自動制御連合講演会, 2006 年 11 月 26 日, 神戸大学.
  23. Kazunori Iwata and Akira Hayashi, A Dissimilarity Measure Based on the Same Principle as Ward's Measure in Hierarchical Clustering Evaluation, 第 9 回情報論的学習理論ワークショップ, 2006 年 11 月 1 日, 大阪大学 中之島センター.
  24. Kazunori Iwata and Akira Hayashi, Theory of a Probabilistic-Dependence Measure of Dissimilarity among Multiple Clusters, The 16th International Conference on Artificial Neural Networks, 12/9/2006, Athens, Greece.
  25. 斧城 悠大, 岩田 一貴, 林 朗, 階層的クラスタリングにおけるWard法と情報理論的な非類似度との関係, 第 5 回情報科学技術フォーラム, 2006 年 9 月 5 日, 福岡大学.

26. 岩田 一貴, 池田 和司, 酒井 英昭, マルコフ決定過程に基づくマルチエージェントシステムの漸近的性質, 第5回情報科学技術フォーラム, 2006年9月7日, 福岡大学.

[その他]

ホームページ等

<http://www.prl.info.hiroshima-cu.ac.jp/~kiwata/>

## 6. 研究組織

### (1) 研究代表者

岩田 一貴 (IWATA KAZUNORI)

広島市立大学・大学院情報科学研究科・助教  
研究者番号：20405492

### (2) 研究分担者

なし

### (3) 連携研究者

なし