

科学研究費助成事業 研究成果報告書

平成 28 年 5 月 27 日現在

機関番号：12601

研究種目：新学術領域研究（研究領域提案型）

研究期間：2011～2015

課題番号：23120004

研究課題名（和文）予測と意思決定のための機械学習理論の構築とその神経回路での実現

研究課題名（英文）Development of Machine Learning Theory for Prediction and Decision Making and Its Realization in Neural Networks

研究代表者

杉山 将（Sugiyama, Masashi）

東京大学・新領域創成科学研究科・教授

研究者番号：90334515

交付決定額（研究期間全体）：（直接経費） 74,300,000円

研究成果の概要（和文）：近年、機械学習に基づく知的情報処理技術が、科学・産業の様々な分野で活用されている。しかし、最先端の機械学習技術をもってしても、強い非線形性・非定常性を持つ高次元かつ複雑なデータからロバストに学習を行なうことは困難である。そこで本研究プロジェクトでは、特徴選択・特徴抽出・強化学習のための新しい機械学習アルゴリズムを開発し、その有効性を高自由度ヒューマノイドロボットの制御を通して実証した。

研究成果の概要（英文）：Recently, intelligent data analysis technology based on machine learning has been used in various areas of science and industry. However, even with the state-of-the-art machine learning methods, it is difficult to learn robustly from high-dimensional complex data having strong non-linearity and non-stationarity. In this research project, we developed novel machine learning algorithms for feature selection, feature extraction and reinforcement learning, and demonstrated their usefulness through high-degree-of-freedom humanoid robot control.

研究分野：知能情報学

キーワード：予測 意志決定 機械学習 特徴選択 強化学習

1. 研究開始当初の背景

機械学習の分野では、情報科学の立場から数多くの学習アルゴリズムが開発されており、その有用性が様々な応用分野で実証されつつある。しかしながら最先端の機械学習技術をもってしても、強い非線形性・非定常性を持つ高次元かつ複雑なデータからロバストに学習を行なうことは未だ困難である。一方ヒトの脳内では、神経回路のように非定常かつノイズの影響が深刻な系においても、ロバストな意志決定システムの構築が行われている。

2. 研究の目的

このように、情報科学的な機械学習の技術とヒトの学習の間には、大きなギャップが存在する。そこで本研究では、情報科学的な機械学習研究を神経科学の知見と有機的に融合させることにより、予測と意思決定の新しい機械学習理論を構築するとともに、実用的な機械学習アルゴリズムを開発する。そして、その有効性をロボット制御などの実世界応用を通して実証する。

3. 研究の方法

(1) 特徴選択・特徴抽出：ヒトの脳内計算機構のようにロバストな予測を実現するためには、冗長な情報源から予測に必要な特徴を正確かつロバストに選択・抽出する必要がある。そこで以下の3つの課題に取り組む。

①特徴選択では、どの入力特徴が出力値と強く関連しているかを正確に見積もることが重要である。そこで本研究では、入力特徴と出力値との間の統計的依存性を見積もるための情報量推定技術を開発する。古典的には特徴量や出力値の確率分布を推定することによって統計的依存性を評価していたが、一定の精度を担保するためには、多数のデータが必要となる。そこで、確率分布を推定せずに統計的依存性を直接的に評価する新しい方法論を構築する。

②特徴の組合せ数は特徴数に対して指数関数的に増加するため、効率のよい探索技術が必要となる。このような組合せ爆発の問題は、解をスパースにする効果を持つ正則化技術を活用することによって効果的に回避でき、線形回帰問題に対する特徴選択を効率よく行える。そこで本研究では、このスパース正則化を応用した新しい特徴探索技術を開発し、特徴選択の精度と計算のスケラビリティを両立させる。

③特徴選択と並んで重要な課題である特徴抽出では、特徴の部分集合でなく低次元表現を求める。一部の基礎技術は特徴選択と共通するため、特徴抽出に関しても並行して研究開発を進める。

(2) 強化学習：強化学習では得られる報酬の期待値が最大になるように制御機構を最適化するが、神経回路のように非定常かつノイズの影響が深刻な系においてはロバスト

性が非常に重要となる。そこで本研究では、①状態遷移のモデル化を含まないモデルフリー型、②状態遷移のモデル化を含むモデルベース型、のそれぞれの観点から高次元制御問題に対する強化学習アルゴリズムを開発する。そして、③開発した強化学習手法を高自由度ヒューマノイドロボットの制御に応用し、その有効性を実証する。

4. 研究成果

(1) 特徴選択・特徴抽出

①特徴量の評価のため、同時確率分布と周辺確率分布積の比および差を、個別に確率分布を推定することなく直接的に推定する手法を開発した。これらの手法は、大域的な最適解を解析的かつ効率よく求められるため、大規模なデータに対して容易に適用できる。

更に、確率密度関数の微分情報が特徴の評価に重要な働きをすることを見出し、確率密度関数を推定することなくその微分を直接的に推定する手法を開発した。この手法も、大域的な最適解を解析的かつ効率よく求められるため、大規模なデータに対しても容易に適用できる。

また、こうした特徴量の有用性評価の技術は、特徴選択だけでなく、特徴抽出やクラスタリングにも自然に応用できることも見出し、その有効性を実験的に示した。

②従来の線形回帰に対するスパース正則化技術は、線形の入出力依存関係しか評価できなかった。そこで、情報量推定技術とスパース正則化技術を組み合わせた新しい非線形特徴選択技術を開発した。また、再生核ヒルベルト空間埋め込みに基づく情報量推定技術を用い、数万もの高次元の特徴から実用的な時間で特徴選択を行える高速アルゴリズムも開発した。

線形回帰よりも高い予測精度を誇るカーネル法においては、カーネル関数を適切に選択することが重要である。古典的な線形回帰のスパース正則化技術を無限次元の関数空間に拡張することによって、複数のカーネルから適切な組合せを見つけられることを見出した。この考え方に従い、効率のよいカーネル組合せ学習アルゴリズムを開発するとともに、その理論的性質を明らかにした。また、同様な正則化技術を条件付き確率密度推定における特徴選択にも応用できることも見出し、アルゴリズムを開発した。

③特徴抽出に関しては、最近傍識別器に対する特徴重み学習アルゴリズム、弱ラベルモデルのエントロピー最大化に基づく計量学習アルゴリズム、KL距離推定のための計量学習アルゴリズムを開発した。更に、確率密度関数そのものを推定せずに確率密度関数の微分を推定できるアルゴリズムを用いて、高精度かつ計算効率のよい教師付き特徴抽出アルゴリズムを開発した。

同時確率分布や周辺確率分布を推定することなく、条件付き確率分布を直接的に推定

できる LSCDE とよばれるアルゴリズムに次元削減を組み合わせたとき、次元削減を行ってから条件付き確率分布を推定するのではなく、次元削減と条件付き確率分布の推定を同時に実行できるアルゴリズムを開発し、推定精度を改善できることを示した。

変分ベイズ推論の枠組みにおいて、確率的主成分分析・確率的行列分解アルゴリズムにおける大域的最適解を解析的に導出するとともに、その性質を理論的に明らかにした。更に、スパース加法モデルに基づく計算効率のよい変分ベイズ推論アルゴリズムを導出し、その有効性を計算機実験により実証した。

データがベクトルでなく固有の構造を持つとき、その構造を活かすことにより学習精度が改善できると期待される。そこで、テンソル構造を持つデータの学習にスパース正則化技術を応用し、実用的な回帰・分類アルゴリズムを開発するとともに、その理論的性質を明らかにした。また、正定値性を持つ行列データに対して、特徴空間の幾何構造を活かした主成分分析アルゴリズムを開発した。

(2) 強化学習

①PGPE とよばれるモデルフリー型の直接政策探索方式の強化学習アルゴリズムの理論解析を行ない、標準的な設定のもとで従来の手法よりも政策勾配の分散を小さくできることを理論的に示すとともに、更に勾配分散を低減させるための最適ベースライン削減法を開発した。そして、より少ないデータから優れた政策を学習できるよう、過去に収集したデータを効果的に再利用する重要度重み付き PGPE アルゴリズムを開発した。それに加え、政策勾配の推定を更に安定化させるため、政策勾配の分散そのものを正則化項として用いるアルゴリズムも開発した。

高次元の状態空間をもつ制御問題では、空間全体において政策を学習することは容易でない。そこで、すでに得られている軌道の周りで政策を学習する新しい強化学習アルゴリズムを構築した。更に、内部モデル学習と軌道モデルベース強化学習を組み合わせたアルゴリズムも開発した。具体的には、動作軌道周りでモデル構築のために、現在の政策をもとに得られたデータから局所的な環境のモデルを効率的な回帰アルゴリズムに基づいて構築した。そして、この局所環境モデルを用いた内部モデルシミュレーションにより、動作周りの軌道改善を行った。これにより、高次元・多自由度の制御対象においてもタスクに合わせた環境モデルを構築でき、その周りで優れた制御政策が獲得できるようになった。

強化学習の研究を進めていくにつれ、強化

学習の最も基本的な構成要素である報酬が、問題によっては自明に定義できないことが明らかになってきた。報酬の定義をデータから学習しようという逆強化学習とよばれる枠組みも議論されているが、必ずしも有効ではない。そこで、政策勾配型のアルゴリズムをオンライン学習の枠組みに拡張し、報酬が任意に変化しても破綻しないオンライン強化学習アルゴリズムを開発した。そして、その頑健性をリグレット最小化という枠組みのもとで理論的に証明した。

②実用的な条件付き確率分布推定法である LSCDE と上記①の PGPE アルゴリズムと組み合わせ、モデルベース型の強化学習法を開発し、ロボットなどの精密なシステムの制御に対して有効であることを示した。更に、上記(1)③で示した LSCDE に特徴抽出を組み込んだ手法をモデルベース PGPE に組み合わせることにより、高次元制御問題に対しても精度良く政策が学習できることを示した。

モデルベース強化学習手法の高次元システムへの適用方法として、モデル予測制御を用いるアプローチを検討した。モデル予測制御は制御政策を柔軟に導出できるというメリットがある一方、ヒューマノイドロボットなどの多自由度・高次元のシステムにおいては、計算時間が膨大になり、実時間制御が困難である。そこで、特異摂動法を援用することにより、対象のロボットモデルを時定数の異なる力学系に分解し、効率的なモデル予測制御によって制御政策を導出する手法を開発した。そして、二次元平面上の歩行ロボットモデルを用いて開発した強化学習法の評価を行い、計算時間の短縮が可能であることを実験的に確認した。

高次元状態空間を持つ強化学習問題において状態空間が因子分解可能な場合、因子間で知見を共有することにより、学習に必要なデータ数を大幅に軽減できることを見出し、マルチタスク型の強化学習アルゴリズムを開発した。

③ヒューマノイドロボットの様な多自由度・高次元系において、上記①②で開発した強化学習手法の有用性を検証した。

軌道モデルベースの強化学習手法を、ヒューマノイドロボットの双腕運動学習課題に適用した。具体的には、ヒューマノイドロボットの両方の手に接続された長い棒の先の動きをロボットの双腕で制御し、その棒先を与えられた経路点を通させるという課題で実験を行なった。実ロボットシステムの双腕での操作課題という複雑な閉リンク系に対して、課題が達成できる制御政策を少ない試行回数で獲得できることを示した。

モデルフリー強化学習の枠組で、先述の過去に収集したデータを効果的に再利用するための重要度重み付き PGPE アルゴリズムをヒューマノイドロボット学習に応用した。具体的には、下図のようなバスケットボールシューティング課題を考え、多自由度・高次元の実システムでも、過去のデータを効率的に活用することにより、複雑な運動課題を達成する制御政策が獲得できることを実験的に示した。



図1：ヒューマノイドの学習により獲得されたバスケットボールシュート動作

これらのロボット制御応用に加え、強化学習の新しい応用として、コンピュータアートに取り組んだ。具体的には、筆をロボットと見立て、筆ロボットに滑らかな線を描画する制御則を強化学習により獲得させる。これにより、写真を毛筆画調に変換することに成功した（ただし、この実験では描画する線の抽出は手作業で行なっている）。

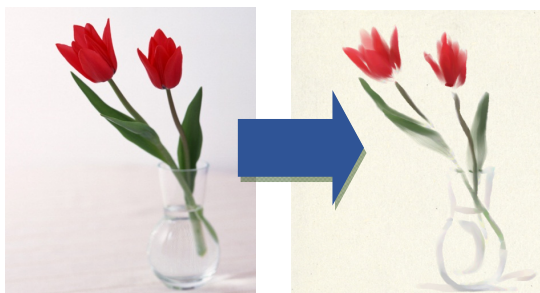


図2：強化学習による毛筆画生成

更に、上記の毛筆画生成における報酬を逆強化学習によって学習するシステムも構築し、パラメータの設定コストを軽減した。



図3：逆強化学習による報酬の学習

5. 主な発表論文等

〔雑誌論文〕(計35件)

1. Sugimoto N, Tangkaratt V, Wensveen T, Zhao T, Sugiyama M, Morimoto J, (2016). Trial and error: using previous experiences as simulation models in humanoid motor learning. IEEE Robotics and Automation Magazine, Vol.23, Issue 1, pp.96-105. (査読あり)
2. Gams A, Petric T, Do M, Nemeč B, Morimoto J, Asfour T, Ude A. (2016). Adaptation and coaching of periodic motion primitives through physical and visual interaction. Robotics and Autonomous Systems, Vol.75, Part B, pp.340-351. (査読あり)
3. Tangkaratt V, Xie N, Sugiyama M. (2015). Conditional density estimation with dimensionality reduction via squared-loss conditional entropy minimization. Neural Computation, vol.27, no.1, pp.228-254. (査読あり)
4. Matsubara T, Uchikata A, Morimoto J. (2015). Spatiotemporal synchronization of biped walking patterns with multiple external inputs by style-phase adaptation. Biological Cybernetics (Web), Vol.109, Issue 6, pp.597-610. (査読あり)
5. Yamada M, Jitkrittum W, Sigal L, Xing E P, Sugiyama M. (2014). High-dimensional feature selection by feature-wise kernelized lasso. Neural Computation, vol.26, no.1, pp.185-207. (査読あり)
6. Lisi G, Noda T, Morimoto J. (2014). Decoding the ERD/ERS: influence of afferent input induced by a leg assistive robot. Frontiers in Neuroscience, Vol.8, Article85, pp.1-12. (査読あり)

7. Niu G, Dai B, Yamada M, Sugiyama M. (2014). Information-theoretic semi-supervised metric learning via entropy regularization. *Neural Computation*, vol. 26, no. 8, pp. 1717-1762. (査読あり)
8. Nakajima S, Sugiyama M., Babacan D, Tomioka R. (2013). Global analytic solution of fully-observed variational Bayesian matrix factorization. *Journal of Machine Learning Research*, vol. 14 (Jan.), pp. 1-37. (査読あり)
9. Schiebener D, Morimoto J., Asfour T, Ude A. (2013). Integrating visual perception and manipulation for autonomous learning of object representations. *Adaptive Behavior*, Vol. 21, No. 5, pp. 328-345. (査読あり)
10. Zhao T, Hachiya H, Tangkaratt V, Morimoto J., Sugiyama M. (2013). Efficient sample reuse in policy gradients with parameter-based exploration. *Neural Computation*, vol. 25, no. 6, pp. 1512-1547. (査読あり)

[学会発表] (計 4 5 件)

1. Ugurlu B, Forni P, Doppmann C, Morimoto J. (2015). Torque and variable stiffness control for antagonistically driven pneumatic muscle actuators via a stable force feedback controller, *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS2015)*, pp. 1633-1639, Hamburg (Germany), Sep. 28-Oct. 2, 2015. (査読あり)
2. Furukawa J, Noda T, Teramae T, Morimoto J. (2015). Estimating joint movements from observed EMG signals with multiple electrodes under sensor failure situations towards safe assistive robot control, *IEEE International Conference on Robotics and Automation (ICRA2015)*, pp. 4985-4991, Seattle (USA), May 26-30, 2015. (査読あり)
3. Sasaki H, Hyvarinen A, Sugiyama M. (2014). Clustering via mode seeking by direct estimation of the gradient of a log-density. In *Machine Learning and Knowledge Discovery in Databases Part III Lecture Notes in Computer Science*, vol. 8725, pp. 19-34, Nancy (France), Sep. 15-19, 2014. (査読あり)
4. Ariki Y, Inamura T, Morimoto J. (2014). Observing human movements to construct a humanoid interface, *IEEE/RAS International Conference on Humanoid Robots (Humanoids2014)*, Madrid (Spain), Nov. 18-20, 2014, pp. 342-347. (査読あり)
5. Vuga R, Ogrinc M, Gams A, Petric T, Sugimoto N, Ude A, Morimoto J. (2013). Motion capture and reinforcement learning of dynamically stable humanoid movement primitives, *IEEE International Conference on Robotics and Automation (ICRA2013)*, Karlsruhe (Germany), May 6-10, 2013, pp. 5264-5270. (査読あり)
6. Xie N, Hachiya H, Sugiyama M. (2012). Artist agent: A reinforcement learning approach to automatic stroke generation in oriental ink painting. In *Proceedings of 29th International Conference on Machine Learning (ICML2012)*, Edinburgh (Scotland), Jun. 26-Jul. 1, 2012, pp. 153-160. (査読あり)
7. Suzuki T, Sugiyama M. (2012). Fast learning rate of multiple kernel learning: Trade-off between sparsity and smoothness. In *Proceedings of Fifteenth International Conference on Artificial Intelligence and Statistics (AISTATS2012) JMLR Workshop and Conference Proceedings*, vol. 22, La Palma (Spain), Apr. 21-23, 2012, pp. 1152-1183. (査読あり)
8. Zhao T, Hachiya H, Niu G, Sugiyama M. (2011). Analysis and improvement of policy gradient estimation. In *Advances in Neural Information Processing Systems 24*, Granada (Spain), Dec. 12-17, 2011, pp. 262-270. (査読あり)

[図書] (計 7 件)

1. 井手 剛, 杉山 将. (2015). 異常検知と変化検知. 192 ページ, 講談社 東京.
2. 杉山 将. (2015). 機械学習のための確率と統計. 128 ページ, 講談社 東京.
3. Sugiyama M. (2015). *Introduction to Statistical Machine Learning*. 534 pages, Morgan Kaufmann.
4. Sugiyama M. (2015). *Statistical Reinforcement Learning: Modern Machine Learning Approaches*. 206 pages Chapman and Hall/CRC.

5. 杉山 将. (2013). イラストで学ぶ機械学習：最小二乗法による識別モデル学習を中心に. 232 ページ, 講談社.

[その他]

ホームページ

○杉山将

<http://www.ms.k.u-tokyo.ac.jp/sugi/>

○森本淳

<http://www.cns.atr.jp/~xmorimo/>

6. 研究組織

(1) 研究代表者

杉山 将 (SUGIYAMA, Masashi)

東京大学・大学院新領域創成科学研究科・教授

研究者番号：90334515

(2) 研究分担者

森本 淳 (MORIMOTO, Jun)

株式会社国際電気通信基礎技術研究所

・脳情報通信総合研究所・研究室長

研究者番号：10505986