

# 位相空間のアフィン変換による 物理系の適応的モデリングと機械学習への応用

中村 真吾<sup>†</sup>橋本 周司<sup>‡</sup>早稲田大学理工学研究科<sup>†</sup>早稲田大学理工学術院<sup>‡</sup>

## 1. はじめに

これまで、機械学習を用いた実機ロボットの最適制御の研究はいくつか提案されている<sup>[1]</sup>。しかし、多大な探索を必要とする問題に対して適用することは、膨大な学習コストを要するため難しくなる。筆者らはこれまで振り振上げ問題を取り上げ、実機のシミュレータを実際の動作データから直接作成し、計算機上で目的の動作を学習させる手法を提案してきた<sup>[2]</sup>。これにより、シミュレータの物理モデルを逐一構築する手間を省き、実機への負荷を軽減し、目的の動作を得る事ができる。

しかし、機械系が変化し、実機の振る舞いが変化した場合には、再びシミュレータの構築とコントローラの学習をし直す必要がある。このように変化のある度に、始めから学習を行うのは効率的ではない。そこで、本研究では目的の力学系の初期学習後、シミュレータの生成する位相空間での振る舞いにアフィン変換を適用し、シミュレータの再モデリングを行うことを検討した。これにより、実機から多量の動作データを再収集する必要がなくなる。ここでは、本手法を振子の振り上げ問題(図3)に適用し、強化学習を用いた適応的な制御の実験結果を報告する。

## 2. システム構成

本研究におけるシステムは、ハードウェア(実機)、バッファ、シミュレータ、コントローラの4つのモジュールからなる(図1)。

実機は、モータと振子から構成され(図2)、振り上げにより倒立させることが目標である。そのために、モータドライバに与える命令電圧  $V$  を制御することでモータの回転軸に発生するトルク  $\tau$  を制御する。また、モータに取り付けられたパルスエンコーダにより、角度  $\theta$  及び角速度  $\omega$  を観測できる。

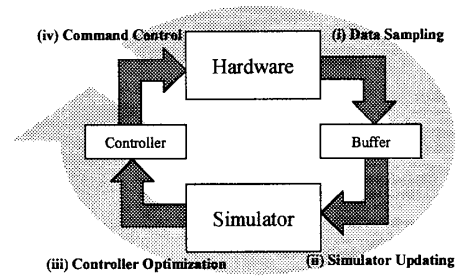


図1. システム構成

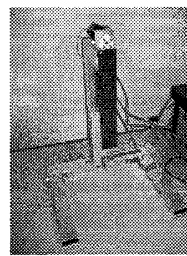


図2. 実機の振子

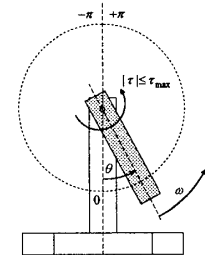


図3. 振り上げ問題

バッファは実際の実機の動作データを保持し、シミュレータの構築のためにデータを使用する。

実機の力学系は、3層ニューラルネットワークによってシミュレートされる。ニューラルネットワークは、時刻  $t$  における振子の角度  $\theta$  と角速度  $\omega$  と D/A ボードからの命令電圧  $V$  を入力パラメータとし、1ステップ後の角度と角速度の増加量である  $\Delta\theta$  と  $\Delta\omega$  を出力する。

コントローラはシミュレータのみを使って目的の振り上げ動作を学習する。ここでは、Actor-Critic アーキテクチャを使った強化学習手法を使用する。

## 3. 提案手法

扱う物理系が現在の状態のみに依存するようなシステムである場合、その振る舞いは位相空間上で一意に記述することができる。実機が変化した時、その振る舞いの様相は変化するが、根本的な構造が変化しない限り、非線形性の強い大きな様相の変化は生じない。そこで、変化前の位相空間での振る舞いに、アフィン変換を適用

Adaptive modeling of physical system with affine transform of phase space for machine learning

<sup>†</sup> School of Science and Engineering, Waseda University

<sup>‡</sup> Faculty of Science and Engineering, Waseda University

し、変化後の振る舞いを近似することを考える。

ニューラルネットワークにより表現されるシミュレータは位相空間での振る舞いを表現している。そこで入力 $(\theta, \omega, V)$ に対する変化前の出力 $(\Delta\theta_1, \Delta\omega_1)$ に対し、アフィン変換を適用したものを変化後のシミュレータの出力 $(\Delta\theta_2, \Delta\omega_2)$ とする。しかし、位相空間内での振舞いは、命令電圧 $V$ に依存して異なるため、変換されるパラメータに $V$ を追加する。この場合、変換式は

$$\begin{pmatrix} \Delta\theta_2 \\ \Delta\omega_2 \\ V \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \Delta\theta_1 \\ \Delta\omega_1 \\ V \end{pmatrix} \quad (1)$$

と表される。ここで $a_{ij}$ はアフィン変換行列の要素である。アフィン変換は、変化前のシミュレータの出力を参照し、変化後に実機から得られるデータをもとに、最小二乗法を使って行列要素を算出する。これにより新たに得られたシミュレータを用いて、計算機上で再びコントローラの最適化を行い、変化後の実機に対する新たなコントローラを取得する。このように変化前に蓄積された情報を利用する事で、まったくの初期状態から再学習を行う必要がなく、実機への負荷を軽減することができる。

#### 4. 実験

実験では、振子の長さが変化することを想定して、図4と表1に示す長さの異なる2つの振子を使用した。始めに、短い振子(以降、振子1)を用い、2章で述べた4つのモジュールを初期化し、振子1のシミュレータ及び、振上げ動作を行うコントローラを取得した。その後、長い振子(以降、振子2)に付け替え、力学系を変化させた。アフィン変換行列を求めるために必要なデータは、振子1に最適化された制御器を用いて振子を制御し、自動でバッファに収集した。バッファからランダムに1000点のデータ(10秒間)を選び、アフィン変換行列を求めた。式(1)により変換されたシミュレータを用い、強化学習を使って振子2の最適なコントローラを取得し、実機に適用し、実際に制御を行った。

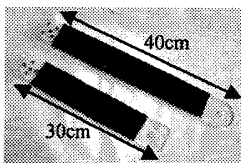


図4. 2種類の振子

表1. 慣性モーメント

Pendulum	Inertia moment (kg·m <sup>2</sup> )
Long	$5.8 \times 10^{-3}$
Short	$2.4 \times 10^{-3}$

#### 5. 結果

図5に、振子1のシミュレータと振子2のデータからアフィン変換されたシミュレータが描く位相空間上の軌跡を示す。慣性モーメントの大きい振子2は、摩擦の影響が少なく、減衰するのに時間がかかることが再現されている。

図6に、アフィン変換前後のシミュレータにより得られたコントローラによる、実機への適用結果を示す。アフィン変換を行うことにより、振り上げの回数が減少し、振子の長さの変化に適應していることが見られる。

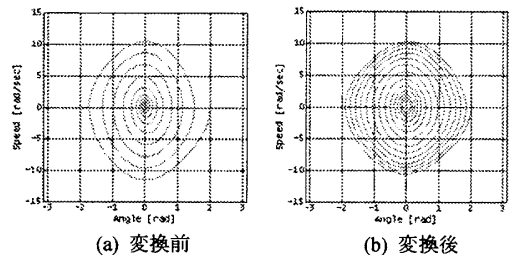


図5. アフィン変換適用前後における位相空間上の軌跡

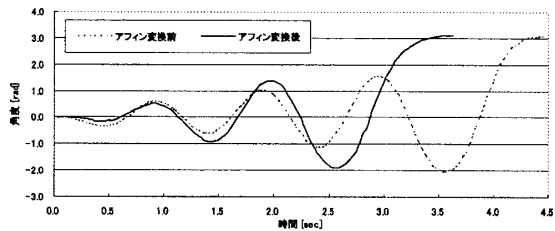


図6. アフィン変換適用前後の制御結果

#### 6. おわりに

本研究では、振子の振上げ問題に対し、シミュレータにアフィン変換を適用することで、変化した実機の再モデリングを行った。再モデリングされたシミュレータにより得られたコントローラは、変化に適應した制御を行うことを確認した。今後はより複雑な機構への応用を考えている。

#### 参考文献

- [1] Bongard J., Zykov V., Lipson H., "Resilient Machines Through Continuous Self-Modeling", Science Vol. 314. no. 5802, pp. 1118-1121, 2006.
- [2] Shingo NAKAMURA, Ryo SAEGUSA, Shuji HASHIMOTO, "Hybrid Learning Strategy for Real Hardware of Swing-up Pendulum", JACIII, Vol.11, No.8, pp.972-978, 2007.