ハミルトン系の変分対称性に基づく学習最適制御*

藤本 健治[†]・堀内 哲[‡]・杉江 俊治[§]

Iterative Learning Optimal Control of Hamiltonian Systems Based on Variational Symmetry*

Kenji FUJIMOTO[†], Tetsu HORIUCHI[‡] and Toshiharu SUGIE[§]

This paper proposes a novel iterative learning control method for Hamiltonian control systems which can solve a class of optimal control problems. First of all, a symmetric property of the input-output mappings of Hamiltonian systems is clarified which plays an important role in solving optimal control problems by gradient method. A concrete learning algorithm is derived for mechanical systems possibly with input saturation. Furthermore, numerical simulations of a 2 link robot manipulator demonstrate the effectiveness of the proposed method.

1. はじめに

ある評価関数を最小にする入力を求める最適制御問題は、制御工学において古くから取り扱われている重要な問題である [1]。最適制御問題を解くためには制御対象の正確で詳細な情報が必要不可欠であり、とくにフィードフォワード入力を求める制御問題では、モデル化誤差による影響を受けやすい点が問題となっていた。本論文はこのような最適制御問題を実験の繰返しによって、学習的に解く新しい手法を提案するものである。これによって制御対象の詳細なモデルを必要とせずに、最適制御手法を用いることができる。

学習制御とは未知の制御対象や信号を推定しつつ制御を行う方法をさす。制御工学の分野以外にもシステム工学の分野などで、さかんに研究が行われているテーマである。制御工学における学習制御とは、フィードバック補償器などの有限個のパラメータチューニングやそれに類するものと、無限次元の空間を探索してフィードフォ

ワード入力を求めるものの二つに大別される [2–4]。どちらも評価関数を作ってその勾配を推定することで、勾配法による最適化によって学習を行っているが、前者のパラメータチューニングの場合には、有限個の実験で比較的容易にその勾配が推定できることが知られている。しかし、後者の問題では無限次元の空間であるフィードフォワード入力を探索する問題となり、一般には解くことが非常に困難な問題となる。上に述べた最適制御問題もこの種類の問題に分類される。有本ら [4] は、制御対象が直達項をもつことを仮定し、その直達項の情報を用いることで軌道追従問題に対するフィードフォワード入力が導出できることを示した。その後この方法は多くの研究者によって発展させられている。

筆者らは、同じく軌道追従問題に対する学習制御手法を提案している [5] が、これは制御対象の直達項の情報をを用いるのではなく、制御対象をハミルトン系に限定することで、その対称性を用いるものであった。この方法は従来法に比べて追従誤差が単調減少するなどの特徴を有している。本論文ではこの結果を拡張し、軌道追従問題だけではなく、より一般的な最適制御問題を学習的に解くアルゴリズムを提案するものである。結果として最適なフィードフォワード入力を求める問題を、有限次元近似することなく直接扱えるという他にない性質をもつものとなっている。

本論文の構成は以下の通りである。2 節ではハミルトン系の変分対称性と、それに付随するいくつかの性質を示す。3 節では対称性を用いて最適制御をいかに学習に

* 原稿受付 2007 年 4 月 23 日

[†] 名古屋大学 大学院 工学研究科 機械理工学専攻 Department of Mechanical Science and Engineering, Graduate School of Engineering, Nagoya University

[‡] (株) 本田技術研究所 二輪開発センター Honda R&D Co., Ltd., Motorcycle R&D Center

[§] 京都大学 大学院 情報学研究科 システム科学専攻 Department of Systems Science, Graduate School of Informatics, Kyoto University

Key Words: iterative learning control, optimal control, Hamiltonian systems, electro-mechanical systems.

よって解くかの原理を説明する．4節では機械系を対象とした場合の具体的な設計手順とアルゴリズムを与える．5節では，提案法を2リンクマニピュレータに適用し，シミュレーションによってその有効性を示す．最後に6節でまとめる．

2. ハミルトン制御系の変分系と対称性

まず，本論文の基礎となるハミルトン制御系の変分系とその対称性について説明を加えておく．ここではハミルトンの正準方程式を一般化したような次式で表される制御系を考える．

$$\Sigma: \begin{cases} \dot{x} = (J-R) \frac{\partial H(x,u,t)}{\partial x}^T, & x(t^0) = x^0 \\ y = -\frac{\partial H(x,u,t)}{\partial u}^T \\ x^1 = x(t^1) \end{cases} \quad (1)$$

ここで， $u(t), y(t) \in \mathbb{R}^m, x(t) \in \mathbb{R}^n$ は入力・出力・状態を表す．ただし，正方行列 $J, R \in \mathbb{R}^{n \times n}$ はそれぞれ歪対称 $J = -J^T$ および対称半正定 $R = R^T \geq 0$ であり，構造行列・散逸行列とよばれる．スカラー関数 $H(x,u,t) \in \mathbb{R}$ はハミルトン関数とよばれ，系の力学的エネルギーおよび入力によって作られる仮想的なポテンシャルを表す．有限の時間区間 $\Omega := [t^0, t^1]$ 上でのこの制御系の挙動を考える．この系において，初期状態 x^0 と入力 $u \in L_2(\Omega)$ を決めると， Ω 上での系の挙動が決まるので，この系を $\Sigma: (x^0, u) \mapsto (x^1, y)$ の作用素 ($\Sigma: \mathbb{R}^n \times L_2^m(\Omega) \rightarrow \mathbb{R}^n \times L_2^m(\Omega)$) ととらえる．また初期状態 x^0 を固定したときの入出力 $u \mapsto y$ の作用素を $\Sigma^{x^0}: L_2^m(\Omega) \rightarrow L_2^m(\Omega)$ と書くことにする．

Σ の変分系 [6,7] とは，作用素としてとらえたときの Fréchet 微分を表す系である．作用素 $\Sigma: u \mapsto y$ の Fréchet 微分とは，各 u に対して

$$\Sigma(u+v) - \Sigma(u) = d\Sigma(u)(v) + o(\|v\|)$$

を満たす， v に関して線形な作用素 $d\Sigma(u)(v)$ である．ハミルトン系 Σ の変分系 $d\Sigma$ は，次のような性質を持つことが知られている．

【補題 1】 [5] $\Sigma(x^0, u)$ の微分 $d\Sigma(x^0, u)(\cdot)$ は任意の (x^0, u) に対してハミルトン系であり，さらに

$$NJ = -JN, \quad NR = RN \quad (2)$$

$$\begin{pmatrix} N & 0 \\ 0 & I \end{pmatrix} \frac{\partial^2 H(x,u,t)}{\partial(x,u)^2} = \frac{\partial^2 H(x,u,t)}{\partial(x,u)^2} \begin{pmatrix} N & 0 \\ 0 & I \end{pmatrix} \quad (3)$$

を満たす行列 $N \in \mathbb{R}^{n \times n}$ が存在するとき， $(d\Sigma(\cdot))^*$ は， $d\Sigma(\cdot)$ を時間反転した状態空間表現をもつ．それらは以下で与えられる．

$$d\Sigma(x^0, u) : (x_v^0, u_v) \mapsto (x_v^1, y_v)$$

$$\begin{cases} \dot{x} = (J-R) \frac{\partial H(x,u,t)}{\partial x}^T, & x(t^0) = x^0 \\ \dot{x}_v = (J-R) \frac{\partial H_v(x_v, u_v, x, u, t)}{\partial x_v}^T, & x_v(t^0) = x_v^0 \\ y_v = -\frac{\partial H_v(x_v, u_v, x, u, t)}{\partial u_v}^T \\ x_v^1 = x_v(t^1) \end{cases} \quad (4)$$

$$(d\Sigma(x^0, u))^* : (x_a^1, u_a) \mapsto (x_a^0, y_a) \begin{cases} \dot{x} = (J-R) \frac{\partial H(x,u,t)}{\partial x}^T, & x(t^0) = x^0 \\ \dot{x}_a = -(J-R) \frac{\partial H_v(x_v, u_a, x, u, t)}{\partial x_v}^T \\ y_a = -\frac{\partial H_v(x_v, u_a, x, u, t)}{\partial u_a}^T \\ x_v(t^1) = -(J-R) N x_a^1 \\ x_a^0 = -N^{-1} (J-R)^{-1} x_v(t^0) \end{cases} \quad (5)$$

ここで変分系 $d\Sigma$ および変分随伴系 $(d\Sigma)^*$ のハミルトン関数 $H_v(x_v, u_v, x, u, t)$ は次式で与えられる．

$$H_v(x_v, u_v, x, u, t) = \frac{1}{2} \begin{pmatrix} x_v \\ u_v \end{pmatrix}^T \frac{\partial^2 H(x, u, t)}{\partial(x, u)^2} \begin{pmatrix} x_v \\ u_v \end{pmatrix}$$

ここでハミルトン系の変分系の入力・出力・状態をそれぞれ u_v, y_v, x_v で，変分随伴系の入力・出力・状態をそれぞれ u_a, y_a, x_a で表している．この補題の意味することは，条件式 (2) および (3) が成立すれば，変分系 $d\Sigma$ と変分随伴系 $(d\Sigma)^*$ が互いに関係付けられるということである．なお， Σ が線形時不変系のときには，対応する伝達関数行列が対称行列になるので，ここではこの性質を変分対称性とよぶ．いま，区間 Ω の時間反転作用素を

$$(\mathcal{R}(u))(t) := u(t^1 - t), \quad t \in \Omega$$

であわらすと，補題 1 から変分系 $d\Sigma$ とその随伴系 $(d\Sigma)^*$ の間に次のような関係があることがわかる．

【定理 1】 補題 1 の仮定を満たす (1) 式のハミルトン系を考える．二つの入力 $v, w \in L_2^m(\Omega)$ に対して，系の状態の軌跡がそれぞれ $\phi(t), \psi(t) \in \mathbb{R}^n, t \in \Omega$ で与えられ，かつ次式を満たすとする．

$$\mathcal{R} \left(\frac{\partial^2 H(x, u, t)}{\partial(x, u)^2} \Big|_{\substack{x=\phi \\ u=v}} \right) = \frac{\partial^2 H(x, u, t)}{\partial(x, u)^2} \Big|_{\substack{x=\psi \\ u=w}} \quad (6)$$

このとき，次の関係が成り立つ．

$$\mathcal{S} (d\Sigma(\phi(t^0), v))^* = (d\Sigma(\psi(t^1), w)) \mathcal{S} \quad (7)$$

ただし， $\mathcal{S}: \mathbb{R}^n \times L_2^m(\Omega) \rightarrow \mathbb{R}^n \times L_2^m(\Omega)$ は次式で与えられるものとする．

$$\mathcal{S}(x^0, u) := (-(J-R)N x^0, \mathcal{R}(u))$$

(証明) (7) 式の両辺に $(\xi^0, \zeta) \in \mathbb{R}^n \times L_2^m(\Omega)$ を入力したとして，両辺の出力 $(\xi^1, \eta) \in \mathbb{R}^n \times L_2^m(\Omega)$ を計算して

みよう. まずは (7) 式の右辺は, 補題 1 の (4) 式より,

$$\left\{ \begin{array}{l} \dot{x} = (J-R) \frac{\partial H(x,w,t)}{\partial x}^T, \quad x(t^0) = \psi(t^0) \\ \dot{x}_v = (J-R) \frac{\partial H_v(x_v, u_v, x, w, t)}{\partial x_v}^T \\ u_v = \mathcal{R}(\zeta) \\ \eta = - \frac{\partial H_v(x_v, u_v, x, w, t)}{\partial u_v}^T \\ x_v(t^0) = -(J-R) N \xi^0 \\ \xi^1 = x_v(t^1) \end{array} \right. \quad (8)$$

一方, やはり補題より, (7) 式の左辺の出力は次のように与えられる.

$$\left\{ \begin{array}{l} \dot{x} = (J-R) \frac{\partial H(x,v,t)}{\partial x}^T, \quad x(t^0) = \phi(t^0) \\ \dot{x}_a = -(J-R) \frac{\partial H_v(x_v, \zeta, x, v, t)}{\partial x_v}^T \\ \mathcal{R}(\eta) = - \frac{\partial H_v(x_v, \zeta, x, v, t)}{\partial \zeta}^T \\ x_v(t^1) = -(J-R) N \xi^0 \\ \xi^1 = x_v(t^0) \end{array} \right.$$

この系を区間 Ω 上で時間反転したダイナミクスは, (6) 式を用いると以下のように表せる.

$$\left\{ \begin{array}{l} \dot{x} = (J-R) \frac{\partial H(x,w,t)}{\partial x}^T, \quad x(t^0) = \psi(t^0) \\ \dot{x}_a = (J-R) \left(I, 0 \right) \frac{\partial^2 H(x,w,t)}{\partial (x,w)^2} \begin{pmatrix} x_v \\ u_a \end{pmatrix} \\ u_a = \mathcal{R}(\zeta) \\ \eta = - \frac{\partial H_v(x_v, u_a, x, v, t)}{\partial u_a}^T \\ x_v(t^0) = -(J-R) N \xi^0 \\ \xi^1 = x_v(t^1) \end{array} \right.$$

これは (8) 式と同一のダイナミクスとなることから, 題意が示される. $\square$

この定理によって, ハミルトン系の変分随伴系 $(d\Sigma)^*$ の入出力は, 作用素 $\mathcal{S}$ を用いて変分系 $d\Sigma$ と関連づけられることがわかる.

さらに Σ ではなく, その初期状態を固定した Σ^{x^0} の変分随伴系を考えると, 定理 1 は次のように表される.

【定理 2】 定理 1 において, $v, w \in L_2^m(\Omega)$ および $\phi(t), \psi(t) \in \mathbb{R}^n$ が (6) 式を満たすならば, 次式が成り立つ.

$$(d\Sigma^{\phi(t^0)}(v))^* = \mathcal{R} (d\Sigma^{\psi(t^1)}(w)) \mathcal{R} \quad (9)$$

(証明) $\Sigma^{x^0}(u)$ と $\Sigma(x^0, u)$ の関係は, 射影 $\pi_y: \mathbb{R}^n \times L_2^m(\Omega) \rightarrow L_2^m(\Omega): (x^1, y) \mapsto y$ と, 挿入 $\nu_u: L_2^m(\Omega) \rightarrow \mathbb{R}^n \times L_2^m(\Omega): u \mapsto (0, u)$ を用いて

$$\Sigma^{x^0}(u) = \pi_y \Sigma((x^0, 0) + \nu_u u)$$

と表せる. したがって

$$d\Sigma^{x^0}(u) = \pi_y d\Sigma(x^0, u) \nu_u \quad (10)$$

となる. いま, (7) 式に左から π_y , 右から ν_u を作用させると, (7) 式の右辺は (10) 式より

$$\begin{aligned} \pi_y d\Sigma(\psi(t^1), w) \mathcal{S} \nu_u &= \pi_y d\Sigma(\psi(t^1), w) \nu_u \mathcal{R} \\ &= d\Sigma^{\psi(t^1)}(w) \mathcal{R} \end{aligned}$$

また左辺は,

$$\begin{aligned} \pi_y \mathcal{S}(d\Sigma(\phi(t^0), v))^* \nu_u &= \mathcal{R} \pi_y (d\Sigma(\phi(t^0), v))^* \nu_u \\ &= \mathcal{R} \nu_u^* (d\Sigma(\phi(t^0), v))^* \pi_y^* \\ &= \mathcal{R} (\pi_y d\Sigma(\phi(t^0), v) \nu_u)^* \\ &= \mathcal{R} (d\Sigma^{\phi(t^0)}(v))^* \end{aligned}$$

となる. 第 3 式で (10) の関係式を用いた. したがって

$$\mathcal{R} (d\Sigma^{\phi(t^0)}(v))^* = d\Sigma^{\psi(t^1)}(w) \mathcal{R}$$

であり, $\mathcal{R}^{-1} = \mathcal{R}$ であることから題意を得る. $\square$

この定理で与えられる結果は, 変分随伴系と変分系をより簡潔に結び付けるものであり, 本論文の学習制御において重要な役割を果たすものである.

また補題 1 の仮定 (2) 式および (3) 式が成立する系として, 単純なハミルトン系を与えておく. (1) 式のハミルトン系において, 状態 $x = (q, p) \in \mathbb{R}^{2m}$, ハミルトン関数

$$H(q, p, u) = \frac{1}{2} p^T M(q)^{-1} p + P(q) - u^T q \quad (11)$$

構造行列および散逸行列が

$$J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}, \quad R = \begin{pmatrix} 0 & 0 \\ 0 & -D \end{pmatrix} \quad (12)$$

のように与えられるものを単純なハミルトン系とよぶ. ここで $M(q) \in \mathbb{R}^{m \times m}$ を慣性行列とよび, D は力学系の摩擦項を表す.

【命題 1】 (1) 式のハミルトン系が単純なハミルトン系であるとする. 慣性行列 M が q に依存しないならば, この系は補題 1 の仮定 (2) 式および (3) 式を満たす.

(証明) ハミルトン関数のヘッセ行列が

$$\frac{\partial^2 H(q, p, u)}{\partial (q, p, u)^2} = \begin{pmatrix} \partial^2 P(q)/\partial q^2 & 0 & -I \\ 0 & M^{-1} & 0 \\ -I & 0 & 0 \end{pmatrix}$$

であることに注意して, 行列 N を $N = \text{diag}(I, -I)$ と選ぶと, 条件式 (2) および (3) の成立が確認できる. $\square$

この命題が示すように, 本論文で主に扱う機械システムは, 変分対称性もしくはそれに近い性質を持っている. 次節ではこの性質が学習制御にいかん利用できるかを説明する.

3. 学習による最適制御問題の解法

本節では勾配法による最適制御の解法と変分随伴系との関係を示し、前節で導いた対称性に基づいた最適制御の学習的解法を与える。

3.1 最適制御と変分随伴系

まず簡単な場合を取り扱う。ここでも補題1の仮定を満たす(1)式の制御対象を考え、最適制御のための評価関数として、その入出力 $u, y \in L_2^m(\Omega)$ の汎関数 $\tilde{F}(u, y) \geq 0$ をひとつ定める。たとえば次式のようなものである。

$$\tilde{F}(u, y) = \int_{t_0}^{t^1} (\gamma_u(\tau) \|u(\tau)\|^2 + \gamma_y(\tau) \|y(\tau)\|^2) d\tau$$

ここでは、初期状態 $x(t^0) = x^0$ のもとで、この評価関数を最小にするような $u \in L_2^m(\Omega)$ を求める問題を考える。すなわち、実際に最小化する問題は $\Gamma(u) := \tilde{F}(u, \Sigma^{x^0}(u))$ として

$$\min_{u \in L_2^m(\Omega)} \Gamma(u)$$

と表される。この評価関数に対して

$$d\Gamma(u)(du) = \langle \nabla \Gamma(u), du \rangle_{L_2^m(\Omega)} \quad (13)$$

を満たす $\nabla \Gamma(u) \in L_2^m(\Omega)$ を $\Gamma(u)$ の勾配という。勾配法とは、この勾配 $\nabla \Gamma(u)$ を求めて、 $-\nabla \Gamma(u)$ の方向に u を更新していくことで評価関数 $\Gamma(u)$ を減少させる方法である。たとえば正定行列 $K_{(k)} \in \mathbb{R}^{m \times m}$ を用いて

$$u_{(k+1)} = u_{(k)} - K_{(k)} \nabla \Gamma(u_{(k)}), \quad k = 0, 1, 2, \dots \quad (14)$$

のように入力 $u_{(k)}$ を変化させていけばよい。ここで k は学習のステップ数を表す。

実際に $\Gamma(u)$ の勾配を計算してみる。

$$\begin{aligned} d\Gamma(u)(du) &= d\tilde{F}(u, y)(du, d\Sigma^{x^0}(du)) \\ &= d\tilde{F}(u, y) \begin{pmatrix} \text{id} \\ d\Sigma^{x^0}(u) \end{pmatrix} (du) \end{aligned}$$

上式と(13)式より

$$\begin{aligned} \nabla \Gamma(u) &= (d\Gamma(u))^*(1) \\ &= \begin{pmatrix} \text{id} \\ d\Sigma^{x^0}(u) \end{pmatrix}^* (d\tilde{F}(u, y))^*(1) \\ &= \left(\text{id}, (d\Sigma^{x^0}(u))^* \right) \begin{pmatrix} \nabla_u \tilde{F}(u, y) \\ \nabla_y \tilde{F}(u, y) \end{pmatrix} \\ &= \nabla_u \tilde{F}(u, y) + (d\Sigma^{x^0}(u))^* \nabla_y \tilde{F}(u, y) \quad (15) \end{aligned}$$

となる。ここで $\nabla_u \tilde{F}(u, y)$, $\nabla_y \tilde{F}(u, y)$ はそれぞれ評価関数 $\tilde{F}(u, y)$ の u, y に関する(偏)勾配であり、入出力 u, y から直ちに求めることができる。これに比べて制御対象の変分随伴系 $(d\Sigma^{x^0}(u))^*$ の部分は、制御対象(1)式の詳細な情報がなければ求めることのできないものである。逆にこの変分随伴系 $(d\Sigma^{x^0}(u))^*$ が求まれば、直

ちに勾配法を適用できることになる。

3.2 対称性を利用した勾配の推定

ここで制御対象 Σ がハミルトン系であることを思いだそう。定理1の仮定を満たすような入出力が得られるとすれば、(15)式の勾配に現れる変分随伴系 $(d\Sigma^{x^0}(u))^*$ を、定理2の(9)式で置き換えることができる。また(9)式の右辺に現れる変分系 $d\Sigma^{\psi(t^1)}(w)$ は、

$$\begin{aligned} d\Sigma^{\psi(t^1)}(w)(z) &= \lim_{h \rightarrow 0} \frac{\Sigma^{\psi(t^1)}(w+h, z) - \Sigma^{\psi(t^1)}(w)}{h} \\ &\approx \frac{\Sigma^{\psi(t^1)}(w+\epsilon, z) - \Sigma^{\psi(t^1)}(w)}{\epsilon}, \quad |\epsilon| \ll 1 \end{aligned}$$

となり[8]、十分小さな正数 $\epsilon > 0$ を用いて近似できる。上式の右辺に現れる Σ は制御対象の出力そのものであり、2回の実験による出力の差から上式が近似できることがわかる。この手順により勾配が得られれば、勾配法(たとえば(14)式)により、評価関数を減少させるアルゴリズムが導ける。

最適制御問題の一部である軌道追従問題に対しては、従来から一般的な制御対象を扱える反復学習制御法が提案されている[4]。これらのアプローチにおいては、本質的に相対次数0の仮定のもとでの直達項の情報を使って勾配を求めているが、これはまさに上記の変分随伴系の情報を使っていることに相当する。また従来法では軌道追従問題しか扱えないが、本手法ではそれ以外の任意の汎関数で表現される問題を扱うことができる。また文献[9,10]などで扱われている方法では、モデルの情報を既知としてこの変分随伴系を計算して用いている。

3.3 状態を含む評価関数

2.節で与えた制御対象 Σ の入出力は、終端状態 x^1 も含めた形で定式化されているため、これを含めた最適制御問題も取り扱うことができる。評価関数として (x^1, u, y) の関数 $\hat{F}$ を考える。たとえば次のようなものである。

$$\hat{F}(x^1, u, y) = \gamma_{x^1} \|x^1 - x^{1d}\|^2 + \int_{t^0}^{t^1} \gamma_u(\tau) \|u(\tau)\|^2 d\tau \quad (16)$$

ここで $x^{1d} \in \mathbb{R}^n$ は、状態 x が終端時刻 $t = t^1$ で達成すべき所望の終端状態を表す。

この場合でも(15)式と同様の計算を行うと、評価関数の勾配は次のように得られる。

$$\begin{aligned} \nabla \Gamma(u) &= \nabla_u \hat{F}(x^1, u, y) \\ &\quad + (d\Sigma(x^0, u))^* \begin{pmatrix} \nabla_{x^1} \hat{F}(x^1, u, y) \\ \nabla_y \hat{F}(x^1, u, y) \end{pmatrix} \end{aligned}$$

したがって、この場合は定理1を使えば学習による最適制御を行うことができる。ただし、そのためには S を構成するために制御対象の構造散逸行列 $J - R$ および

N の情報が必要となり、また S^{-1} を必要とするため、 $(J-R)$ が正則である必要がある。このように状態を含む評価のためには、より多くの事前情報が必要となることに注意しておく。

3.4 入力飽和の取り扱い

現実の制御対象においては、入力飽和や状態の拘束などといった制約条件が存在する。最適制御問題においても、入力飽和問題は古くから扱われている。いま、外部から加える入力の指令値 $\hat{u}$ に対して、(1) 式の制御対象に働く実効入力 u が

$$u = \theta(\hat{u}) \quad (17)$$

によって表されるものとする。たとえば $u = (u_1, u_2, \dots, u_m)$ のように要素ごとに表現したときに、各要素 u_k がそれぞれ上下限 $u_k^{\max}, -u_k^{\max}$ を持つ場合には、飽和関数 $\theta(\hat{u}) = (\theta_1(\hat{u}_1), \theta_2(\hat{u}_2), \dots, \theta_m(\hat{u}_m))$ の各要素はたとえば次のように表されることになる。

$$\theta_k(\hat{u}_k) = \begin{cases} u_k^{\max} & (\hat{u}_k \geq u_k^{\max}) \\ \hat{u}_k & (|\hat{u}_k| < u_k^{\max}) \\ -u_k^{\max} & (\hat{u}_k \leq -u_k^{\max}) \end{cases}$$

以下では θ が既知で微分可能であると仮定する。もし上式のように微分可能でない場合には、実現可能な制御入力を与えるなめらかな関数で近似するものとする。

このとき制御対象の挙動は、外部入力 $\hat{u}$ を用いて

$$\hat{H}(x, \hat{u}, t) := H(x, \theta(\hat{u}), t)$$

というハミルトン関数を持つハミルトン系となる。この系に対して変分対称性を導く出力 $\hat{y}$ は

$$\begin{aligned} \hat{y} &= -\frac{\partial \hat{H}(x, \hat{u}, t)}{\partial \hat{u}}^T = -\frac{\partial \theta(\hat{u})}{\partial \hat{u}}^T \frac{\partial H(x, u, t)}{\partial u}^T \\ &= \frac{\partial \theta(\hat{u})}{\partial \hat{u}}^T y \end{aligned}$$

のように計算でき、(1) 式のハミルトン系の出力 y に飽和関数のヤコビ行列 $(\partial \theta(\hat{u}) / \partial \hat{u})^T$ をかければよいことがわかる。したがって、飽和関数 θ の情報と入力指令値 $\hat{u}$ がわかれば、入力飽和を持つハミルトン系の学習制御が行えることになる。

4. 機械系の学習最適制御

本節では、前節までに議論したハミルトン系の対称性を用いた学習最適制御の枠組みを機械系に適用し、具体的な学習則を導く。ここでは、(11) 式と (12) 式の構造を有する単純なハミルトン系 (1) 式を対象とする。この系に対しては出力が

$$y = -\frac{\partial H(q, p, u)}{\partial u}^T = q$$

となり、配位座標 q が対称性を導く出力になる。

4.1 一般化正準変換を用いたフィードバック系

3 節で述べたように、本論文では制御対象に加えるフィードフォワード入力を学習によって獲得するものである。しかしこのフィードフォワード入力は、制御対象に対してそのまま入力するのではなく、多くの場合はローカルフィードバックを加えて系を漸近安定化したのち、参照入力としてフィードフォワード入力を与えることが多い。ここでもそのような問題設定を考えるが、その際に閉ループ系がハミルトン系となるようにフィードバック補償器を設計する必要がある。このようなハミルトン系の構造を保存するフィードバック変換を一般化正準変換といい、文献 [11, 12] などにその設計法が与えられている。(11) 式のハミルトン関数を持つ単純なハミルトン系に対しては、正定行列 $K_P, K_D \in \mathbb{R}^{m \times m}$ を用いた

$$u = -K_P q - K_D \dot{q} + \bar{u} \quad (18)$$

のような PD フィードバックがハミルトン系の形を保存する一般化正準変換となることが知られている。変換後のハミルトン系は、以下のようなハミルトン関数 $H^c(q, p)$ 、構造行列 J^c 、散逸行列 R^c を持つ。

$$\begin{aligned} H^c(q, p) &= \frac{1}{2} p^T M(q)^{-1} p + P(q) + \frac{1}{2} q^T K_P q - u^T q \\ J^c &= \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}, \quad R^c = \begin{pmatrix} 0 & 0 \\ 0 & D + K_D \end{pmatrix} \end{aligned}$$

この系は単純なハミルトン系であり、さらに時間軸を適当にとることにより PD ゲイン K_P, K_D を大きくしていけば、その挙動が命題 1 の条件を満たすハミルトン系に近づいていくことが示されている [5]。したがって PD ゲインを十分大きくとれば、3 節のアルゴリズムを用いて、新しい制御入力 $\bar{u}$ に対するフィードフォワード入力の学習を適用できることになる。しかし実用上はゲインを大きくするには限界があり、またハイゲインのフィードバックを施すと最適制御の評価関数の意味が変わってしまうなどの弊害もあることに注意しておく。

4.2 対称軌道の生成

ここでは定理 1 および 2 を用いるための条件式 (6) を満たす軌道 ϕ, ψ の生成法について述べる。ここでの制御対象は命題 1 の条件を満たす単純なハミルトン系であるとする。このとき (6) 式の条件は次のように簡単化される。

$$\mathcal{R} \left(\frac{\partial^2 P(q)}{\partial q^2} \Big|_{\substack{x=\phi \\ u=v}} \right) = \frac{\partial^2 P(q)}{\partial q^2} \Big|_{\substack{x=\psi \\ u=w}}$$

したがって、制御対象が線形系 ($P(q)$ がたかだか 2 次の関数) であれば自動的に満たされ、非線形であっても

$$\mathcal{R} \left(q \Big|_{\substack{x=\phi \\ u=v}} \right) = q \Big|_{\substack{x=\psi \\ u=w}}$$

のように配位座標 q の挙動を時間対称に選べばよいことがわかる．すなわち，ある軌道 $\phi(t)$ が与えられたときに，配位座標 $q(t)$ が時間区間 $\Omega = [t^0, t^1]$ の中心 $t = (t^0 + t^1)/2$ に関して対称になるような実現可能な軌道 $\psi(t)$ を用意してやれば，**定理 1** もしくは **2** によって評価関数の勾配が得られることになる．

制御対象であるハミルトン系 Σ に PD フィードバック (18) 式を施しているため，閉ループ系は Fig. 1 のように表される．図の信号 q および $\dot{q}$ のところに目標となる信号，すなわち与えられた軌道 $\phi(t)$ の時間反転軌道

$$(q^r, \dot{q}^r) = (\mathcal{R}(q|_{x=\phi}), -\mathcal{R}(\dot{q}|_{x=\phi})) \quad (19)$$

を与えることで， q を q^r に追従させて近似的に (6) 式の関係達成する．ここでも PD ゲイン K_P , K_D を大きくすれば，近似の誤差 (追従誤差) は小さくなるため，十分な学習精度が得られない場合にはゲインを大きくしてやればよい．

Fig. 1 Feedback system based on generalized canonical transformations

4.3 学習アルゴリズムの導出

ここでは簡単のため，評価関数 $\tilde{F}(u, y)$ に対して**定理 2**を用いて学習アルゴリズムを導出する．**定理 1**に基づいて， $\hat{F}(x^1, u, y)$ の形の評価関数に対する学習アルゴリズムも同様に導くことができるが，**3.3**節で述べたように制御対象の構造散逸行列 $J-R$ の情報が必要なことや，記述が複雑となることから，ここでは具体的なアルゴリズムの導出は割愛する．

(17) 式で表される入力飽和を持つ系に対する評価関数 $\tilde{F}(\hat{u}, y)$ を考えると，学習アルゴリズムは以下のようになる．ただし括弧添字をつけた $u_{(k)}$, $y_{(k)}$ はそれぞれ k ステップ目の実験の入力・出力などを表すものとする．また機械系に初期速度を任意に与えるのは困難であるため，各実験の初期速度はすべて 0 とした¹．

[学習アルゴリズム]

Step 0 ステップ数を表すカウンタを $k=0$ とし，初期状態 $q_{(0)}^0$ ，入力 $\bar{u}_{(0)}$ ，参照入力 $q_{(0)}^r=0$ を与えて実験を行う．

Step 1 初期状態，入力，参照入力を以下のように定めて $3k+1$ 回目の実験を行う．

$$\begin{aligned} q_{(3k+1)}^0 &= q_{(3k)}^1 \\ \bar{u}_{(3k+1)} &= 0 \\ q_{(3k+1)}^r &= \mathcal{R}(q_{(3k)}) \\ \dot{q}_{(3k+1)}^r &= -\mathcal{R}(\dot{q}_{(3k)}) \end{aligned}$$

Step 2 初期状態，入力，参照入力を以下のように定めて $3k+2$ 回目の実験を行う．

$$\begin{aligned} q_{(3k+2)}^0 &= q_{(3k)}^1 \\ \bar{u}_{(3k+2)} &= \epsilon_{(k)} \mathcal{R} \left(\frac{\partial \theta^T}{\partial \hat{u}} (\hat{u}_{(3k)}) \nabla_y \tilde{F}(\hat{u}_{(3k)}, q_{(3k)}) \right) \\ q_{(3k+2)}^r &= \mathcal{R}(q_{(3k)}) \\ \dot{q}_{(3k+2)}^r &= -\mathcal{R}(\dot{q}_{(3k)}) \end{aligned}$$

Step 3 初期状態，入力，参照入力を以下のように定めて $3k+3$ 回目の実験を行う．

$$\begin{aligned} q_{(3k+3)}^0 &= q^0 \\ \bar{u}_{(3k+3)} &= \bar{u}_{(3k)} - K_{(k)} \times \left(\nabla_{\hat{u}} \tilde{F}(\hat{u}_{(3k)}, q_{(3k)}) \right. \\ &\quad \left. + \frac{1}{\epsilon_{(k)}} \mathcal{R}(q_{(3k+2)} - q_{(3k+1)}) \right) \\ q_{(3k+3)}^r &= \dot{q}_{(3k+3)}^r = 0 \end{aligned}$$

Step 4 k を 1 だけ増やしてステップ 1 へ．

ここで $\epsilon_{(k)} > 0$ は十分小さい正定数， $K_{(k)} > 0$ は勾配法のステップパラメータとなる正定行列で，評価関数が減少する範囲で大きく選ぶ．

このように最適制御問題を解く学習アルゴリズムは，3 回の実験によって勾配法の 1 ステップを実行するアルゴリズムになっている．なお，軌道追従制御問題に対する反復学習 [5] で扱ったように，獲得すべき軌道がそもそも Ω 上で時間対称な場合にはステップ 1 と 3 は同一の実験で代用できるため，2 回の実験で 1 ステップの学習アルゴリズムとなる．

5. 数値例

Fig. 2 2-link robot manipulator

前節で導出した学習最適制御のアルゴリズムを Fig. 2 のような水平面内を運動する 2 リンクのマニピュレータに適用する．この系は関節角 $q = (q_1, q_2)$ ，慣性行列

$$M(q) = \begin{pmatrix} \rho_1 + \rho_2 + 2\rho_3 \cos \theta_2 & \rho_2 + \rho_3 \cos \theta_2 \\ \rho_2 + \rho_3 \cos \theta_2 & \rho_2 \end{pmatrix}$$

$$\rho_1 := I_1 + m_1 l_{g1}^2 + m_2 l_1^2, \quad \rho_2 := I_2 + m_2 l_{g2}^2$$

¹ 非零の初期速度を指定できる場合には，(6) 式の条件がより正確に成立するように (19) 式に基づいて初期速度を与えればよい．

$$\rho_3 := m_2 l_1 l_{g2}$$

および摩擦の行列 $D = \text{diag}(d_1, d_2)$ を用いて (1) 式のハミルトン系で表現できる. 入力 $u = (u_1, u_2)$ は各関節に加えるトルク, m_k はリンク k の質量, l_k はリンク k の長さ, l_{gk} はリンク k の重心から関節までの長さ, I_k はリンク k の重心回りの慣性モーメント, d_k は関節 k の摩擦係数, q_k はリンク k の回転角度を表す. 物理パラメータの具体的な値は, $l_1 = 0.25$, $l_2 = 0.30$, $\rho_1 = 2.55$, $\rho_2 = 0.72$, $\rho_3 = 0.65$, $d_1 = 0.2415$, $d_2 = 0.2457$ とする. またこの系は各軸 $u_1^{\max} = 100$, $u_2^{\max} = 50$ で入力が飽和するものとし, 飽和関数の各要素が以下で与えられるとする.

$$\theta_k(\hat{u}_k) = u_k^{\max} \tanh \frac{\hat{u}_k}{u_k^{\max}} \quad (20)$$

この系に対し, 最適制御の評価関数を以下のように設定した.

$$\tilde{J} := \int_{t^0}^{t^1} \gamma_y(\tau) \|q(\tau) - q^{1^d}\|^2 d\tau$$

$$\gamma_y(t) = \begin{cases} 0 & (t^0 \leq t \leq t^2) \\ \frac{1}{2} \left(1 - \cos \left(\frac{t - t^2}{t^1 - t^2} \pi \right) \right) & (t^2 < t \leq t^1) \end{cases}$$

この $\gamma_y(t)$ は, (16) 式の終端条件 $\|x^1 - x^{1^d}\|^2$ の項を模擬したものである. このように終端条件 x^1 を含まない関数で終端条件に関する評価をすることで, 定理 1 のかわりに定理 2 を用いることができるため, より少ない情報のもとで学習制御が行える. また学習アルゴリズムを適用するためのパラメータを次のように選定した.

$$\Omega = [t^0, t^1] = [0, 0.5], \quad t^2 = 0.45$$

$$q^0 = (0, -\pi/2), \quad q^{1^d} = (\pi/2, \pi/2)$$

$$K_P = 30, \quad K_D = 20, \quad \epsilon(k) = 1$$

まず入力飽和を考えずに次式の学習ゲイン (勾配法のステップパラメータ) を用いて終端状態達成の学習のシミュレーションを行った.

$$K(k) = \text{diag}(10^5, 5.0 \times 10^4)$$

10 ステップの学習実験を行ったときのリンク 2 の関節角 q_2 と入力トルク u_2 の時間応答を Figs. 3, 4 に示す. 点線が学習途中の軌跡であり, 実線が 10 ステップ目の学習後のデータである. 速やかに指定した終端状態を達成しているが, 入力飽和を考慮していないため, その値が大きく上限の 50 [Nm] をこえていることがわかる.

つぎに, (20) 式の入力飽和関数の情報を用いて 3.4 節で導いた入力飽和を考慮した学習アルゴリズムを適用した. 学習ゲインは次のものを用いた.

$$K(k) = \text{diag}(2.5 \times 10^5, 10^5)$$

Fig. 3 Responses of the angle q_2 of link 2 without saturation

Fig. 4 Input u_2 of link 2 without saturation

Fig. 5 Responses of the angle q_2 of link 2 with saturation

Fig. 6 Input u_2 of link 2 with saturation

先の例と同様に 10 ステップの学習実験を行ったときのリンク 2 の関節角 q_2 と入力トルク u_2 の時間応答を Figs. 5,

Fig. 7 History of the cost function J with saturation

6に示す。やはり点線が学習途中の軌跡であり、実線が10ステップ目の学習後のデータである。図を見ると、制約 50 [Nm] 内の入力で指定した終端状態が達成できているのがわかる。その際の評価関数の変化をFig. 7に示す。この図より、少ない回数の学習で速やかに評価関数が減少している様子がわかる。これらのシミュレーション結果により、提案手法の有効性が確認できた。

6. おわりに

本論文ではハミルトン系の変分対称性を利用して、最適制御問題に対するフィードフォワード入力を学習的に解く手法を提案した。機械系に対して具体的なアルゴリズムを導出し、入力飽和を有する系に対して最適なフィードフォワード入力を学習できることをシミュレーションにより確認した。この方法は、これまでの反復学習や反復フィードバックチューニングでは扱えないクラスの問題を解くことのできるアルゴリズムである。実際に、筆者らのグループでは、このアルゴリズムをロボットの最適なホッピング軌道の生成などに応用し一定の成果[13]を得ており、その応用範囲は広いと考えている。

参考文献

- [1] 嘉納：システムの最適理論と最適化，コロナ社 (1987)
- [2] 藤本：無限次元信号の最適化としての学習制御；第50回システム制御情報学会研究発表講演会講演論文集 (2006)
- [3] H. Hjalmarsson: Iterative feedback tuning: an overview; *Int. J. Adaptive Control and Signal Processing*, Vol. 16, pp. 373–395 (2002)
- [4] S. Arimoto, S. Kawamura and F. Miyazaki: Bettering operation of robotics; *Journal of Robotic Systems*, Vol. 1, No. 2, pp. 123–140 (1984)
- [5] K. Fujimoto and T. Sugie: Iterative learning control of Hamiltonian systems: I/O based optimal control approach; *IEEE Trans. Autom. Contr.*, Vol. 48, No. 10, pp. 1756–1761 (2003)
- [6] P. E. Crouch and A. J. van der Schaft: *Variational and Hamiltonian Control Systems*, Vol. 101 of *Lecture Notes on Control and Information Science*, Springer-Verlag (1987)
- [7] P. E. Crouch: Geometric structure in systems theory; *IEEE Proceedings*, Vol. 128, No. 5, pp. 242–252 (1981)
- [8] P. Blanchard and E. Bruning: *Variational Methods in Mathematical Physics: A Unified Approach*, Springer-Verlag (1992)
- [9] M. Yamakita and K. Furuta: Iterative generation of virtual reference for a manipulator; *Robotica*, Vol. 9, pp. 71–80 (1991)
- [10] 十河，足立：線形システムに対する反復学習制御アルゴリズムの設計法；計測自動制御学会論文集，Vol. 31, No. 10, pp. 1601–1607 (2001)
- [11] 藤本：ハミルトニアンシステムの制御；計測と制御，Vol. 39, No. 2, pp. 99–104 (2000)
- [12] A. J. van der Schaft: *L₂-Gain and Passivity Techniques in Nonlinear Control*, Springer-Verlag (2000)
- [13] S. Satoh, K. Fujimoto and S.-H. Hyon: Gait generation for passive running via iterative learning control; *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems* (2006)