

研究速報

直接逆モデリングのオンライン学習特性

大山 英明† (正員) 館 暉††

Learning Characteristics of On-Line Direct Inverse Modeling

Eimei OYAMA†, Member and Susumu TACHI††, Nonmember

† 工業技術院機械技術研究所, つくば市

Mechanical Engineering Laboratory, AIST, Namiki 1-2, Tsukuba-shi, 305 Japan

†† 東京大学工学部, 東京都

Faculty of Engineering, the University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo, 113 Japan

あらまし 直接逆モデリングは代表的な逆モデル学習法である。本論文では、学習要素が1階連続微分可能で、対象システムの入力の次元と出力の次元が等しく、逆モデルの出力する近似解について対象システムのヤコビ行列がフルランクである場合、十分正確な逆モデルが学習された時点において、直接逆モデリングのオンライン学習に目標指向性が存在することを示す。

キーワード 直接逆モデリング, 目標指向性, オンライン学習, バックプロパゲーション学習

1. まえがき

システムの逆モデルを学習するために、対象システムの出力を学習要素に入力し、対象システムの入力を学習要素の教師信号として用いる学習法が、広く用いられてきた。この手法は、Jordanによって直接逆モデリングと名づけられた[1]。この学習の詳細は2.で説明する。対象システムの特定の出力目標値について、学習により逆モデルの精度を改善できるという性質を、学習の目標指向性と言う。直接逆モデリングを利用した多くの研究では、オフライン学習が想定されていたが、その場合学習の目標指向性が存在しないことをJordanは指摘した[1]。一方Albusは、学習した逆モデルを利用して対象を制御しつつ学習を行うオンライン学習を提案していた[2]。Colombanoらはオンライン学習には、目標指向性が存在する場合があることを指摘した[3]。本論文では、(1) 対象システムの入力の次元と出力の次元が等しく、(2) 学習要素が1階連続微分可能で、(3) 逆モデルの出力する近似解について対象システムのヤコビ行列がフルランクである場合、(4) 十分正確な逆モデルが学習された時点において、直接逆モデリングのオンライン学習に目標指向性が存在することを制御論的に示す。

2. 準備

2.1 逆モデル

あるシステムの m 次元入力ベクトルを \mathbf{x} , n 次元

出力ベクトルを \mathbf{y} として、両者の関係を

$$\mathbf{y} = \mathbf{f}(\mathbf{x}) \quad (1)$$

とする。このような対象システム $\mathbf{f}(\mathbf{x})$ に対して、

$$\mathbf{y} = \mathbf{f}(\mathbf{g}(\mathbf{y})) \quad (2)$$

という関係を満たすシステム $\mathbf{g}(\mathbf{y})$ を $\mathbf{f}(\mathbf{x})$ の逆システムと呼ぶ。このような関数 $\mathbf{g}(\mathbf{y})$ があれば、対象システムの出力目標値を \mathbf{y}_d とすると、

$$\mathbf{x}_d = \mathbf{g}(\mathbf{y}_d) \quad (3)$$

により、対象システムへの入力 \mathbf{x}_d を計算することによって、出力 \mathbf{y} を \mathbf{y}_d に一致させることが可能である。 $\mathbf{g}(\mathbf{y})$ を学習要素によって学習することを考える。学習した $\mathbf{g}(\mathbf{y})$ のモデルを対象システムの逆モデルと呼ぶ。

2.2 学習要素

本論文では、パラメータ \mathbf{W} を調節して学習を行う学習要素を想定する。入力が \mathbf{y} であるとき、その出力を $\Phi(\mathbf{W}, \mathbf{y})$ とし、その教師信号を $\Phi'(\mathbf{W}, \mathbf{y})$ とする。 $\Phi(\mathbf{W}, \mathbf{y})$ は、 \mathbf{W} と \mathbf{y} について1階連続微分可能な関数とする。本論文では、学習法として最も広く用いられているバックプロパゲーション学習を想定する[4]。

$$\Phi'(\mathbf{W}, \mathbf{y}) = \mathbf{s} \quad (4)$$

のように、教師信号 \mathbf{s} が与えられた場合、バックプロパゲーション学習では、評価関数

$$S(\mathbf{W}) = \frac{1}{2} |\mathbf{s} - \Phi(\mathbf{W}, \mathbf{y})|^2 \quad (5)$$

が減少するように、

$$\begin{aligned} \frac{d\mathbf{W}}{dt} &= - \left(\frac{\partial S(\mathbf{W})}{\partial \mathbf{W}} \right)^T \\ &= \left(\frac{\partial \Phi(\mathbf{W}, \mathbf{y})}{\partial \mathbf{W}} \right)^T (\mathbf{s} - \Phi(\mathbf{W}, \mathbf{y})) \end{aligned} \quad (6)$$

に従って、 \mathbf{W} を更新する。

2.3 直接逆モデリング

直接逆モデリングでは、学習の時点の対象システムへの入力が \mathbf{x} で、それに対応する出力が \mathbf{y} のとき、

$$\Phi'(\mathbf{W}, \mathbf{y}) = \Phi'(\mathbf{W}, \mathbf{f}(\mathbf{x})) = \mathbf{x} \quad (7)$$

のように教師信号を定める。図1に直接逆モデリングによる学習と制御の概念図を示す。

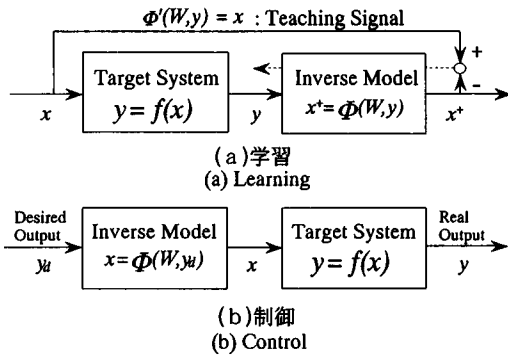


図1 直接逆モデリングの概念図

Fig.1 Conceptual diagram of direct inverse modeling.

学習時の制御対象への入力 x の計算方式には、オフライン学習とオンライン学習の2種類がある。オフライン学習は、出力目標値 y_d と無関係に x を生成して式(7)の学習を行うため、学習の目標指向性が存在しないという問題点が指摘されている[1]。すなわち、逆モデル $\Phi(W, y)$ の精度を、ある特定の目標値 y_d について改善することはできない。

一方、オンライン学習では、目標値 y_d を逆モデルに入力し、

$$x = \Phi(W, y_d) \quad (8)$$

のように対象システムへの入力 x を計算し、それを対象システムに入力したときのシステムの出力

$$y = f(\Phi(W, y_d)) \quad (9)$$

を観測し、次式に従って学習を行う[2]。

$$\begin{aligned} \Phi'(W, y) &= \Phi'(W, f(\Phi(W, y_d))) \\ &= \Phi(W, y_d) \end{aligned} \quad (10)$$

すなわち、図1(b)のように制御を行った後、その結果に基づいて、図1(a)のように学習を行う。オンライン学習には、学習の目標指向性が存在する場合があることが指摘されている[3]。次章でそれを解析する。

3. オンライン学習の目標指向性

時刻 t におけるパラメータ W の値を $W(t)$ とし、出力誤差を

$$e(t) = y_d - f(\Phi(W(t), y_d)) \quad (11)$$

とする。 $t = 0$ の時点で、十分正確な逆モデルが獲得されているものとし、 $e(0)$ は十分小さいものと仮定する。

本論文では学習の目標指向性を

$$\lim_{t \rightarrow \infty} e(t) = 0 \quad (12)$$

が成立することと考え、 y_d が一定という簡単な場合の目標指向性について述べる。以下、 $e(t)$ を e と略す。

$F(e)$ を e について1階連続微分可能な関数とし、 $F(e)$ のヤコビ行列を

$$J(e) = \frac{\partial F(e)}{\partial e} \quad (13)$$

とする。次式のような力学系

$$\frac{de}{dt} = F(e) \quad (14)$$

において、 $e(0)$ が十分小さく、

$$F(0) = 0 \quad (15)$$

が成立し、 $J(0)$ のすべての固有値の実部が負であれば、 $e = 0$ は漸近安定であり、式(12)が成立する。以下、 $\lambda_i(G)$ ($i = 1, 2, \dots, n$) を n 行 n 列の行列 G の固有値とし、 $Re(\lambda_i(G))$ をその実部とする。

3.1 多層神経回路の場合

ここでは、直接逆モデリングのオンライン学習による e の変化を導き、その漸近安定の条件を検討する。学習要素として、バックプロパゲーション学習を行う多層神経回路を想定する。以下、

$$A(x) = \frac{\partial f(x)}{\partial x} \quad (16)$$

$$B(W, y) = \frac{\partial \Phi(W, y)}{\partial y} \quad (17)$$

$$B_w(W, y) = \frac{\partial \Phi(W, y)}{\partial W} \quad (18)$$

とおく。また、 $A(\Phi(W, y_d))$ を A 、 $B(W, y_d)$ を B 、 $B_w(W, y_d)$ を B_w と書くことにする。

バックプロパゲーション学習を行った場合、神経回路のパラメータの変化は

$$\begin{aligned} \frac{dW}{dt} &= B_w(W, f(\Phi(W, y_d)))^T \\ &\quad (\Phi(W, y_d) - \Phi(W, f(\Phi(W, y_d)))) \\ &= B_w(W, f(\Phi(W, y_d)))^T B e + O(e^2) \end{aligned} \quad (19)$$

となる。\$O(e^2)\$ は、各方程式において、ベクトル \$e\$ の 2 次以上の項の総和を示す。

$$B_w(W, f(\Phi(W, y_d))) = B_w(W, y_d - e) \quad (20)$$

であり、

$$\frac{dW}{dt} = B_w^T B e + O(e^2) \quad (21)$$

を得る。まとめると、学習による \$e\$ の変化は、

$$\begin{aligned} \frac{de}{dt} &= -\frac{df(\Phi(W, y_d))}{dt} = -AB_w \frac{dW}{dt} \\ &= -AB_w B_w^T B e + O(e^2) \end{aligned} \quad (22)$$

となる。明らかに式 (15) が成立している。式 (22) の右辺の \$e = 0\$ におけるヤコビ行列は、

$$J(0) = -AB_w B_w^T B \quad (23)$$

であり、以下その固有値を検討する。ここで \$A\$ はフルランクとする。

\$\Phi(W, y)\$ が対象システムの真の逆モデル \$g(y)\$ を近似していることから、

$$\begin{aligned} B(W, y) &= \frac{\partial g(y)}{\partial y} + \Delta B(W, y) \\ &\approx \frac{\partial g(y)}{\partial y} = A^+(y) \end{aligned} \quad (24)$$

が成立する。\$n = m\$ の場合、\$A^+(y)\$ は \$A(\Phi(W, y))\$ の逆行列 \$A^{-1}(\Phi(W, y))\$ である。\$n < m\$ の場合には、\$A(\Phi(W, y))\$ の一般化逆行列のうち、行列 \$\Delta B(W, y)\$ のノルムが最小になるものとする。以下、\$A^+(y_d)\$ を \$A^+\$ と略す。よって、\$A\$ がフルランクで、\$|\Delta B|\$ が十分小さい場合、

$$\begin{aligned} \lambda_i(AB_w B_w^T B) &= \lambda_i(AB_w B_w^T A^+ + AB_w B_w^T \Delta B) \\ &= \lambda_i(AB_w B_w^T A^+) + O_\lambda(\Delta B) \end{aligned} \quad (25)$$

と展開できる。\$O_\lambda(\Delta B)\$ は \$\Delta B\$ の成分の 1 次以上の項の総和を示す。

通常 \$B_w\$ はフルランク (\$rank(B_w) = m\$) とみなすことができる。\$n = m\$ の場合

$$\lambda_i(AB_w B_w^T A^{-1}) = \lambda_i(B_w B_w^T) \quad (26)$$

であり、\$\lambda_i(B_w B_w^T)\$ は正の実数である。\$|O_\lambda(\Delta B)|\$ は逆モデルの精度が向上すれば、いくらでも小さくなるため、十分正確な逆モデルが学習されているという仮定より、

$$Re(\lambda_i(AB_w B_w^T B)) > 0 \quad (27)$$

を得る。これより式 (12) が成立し、目標指向性が存在することがわかる。一方、\$n < m\$ の場合、\$A^+\$ によっては、

$$Re(\lambda_i(AB_w B_w^T A^+)) > 0 \quad (28)$$

が成立しない場合がある。そのため目標指向性が存在するとは限らない。

3.2 ある種の学習要素の場合

学習要素によっては、\$n < m\$ の場合でも、目標指向性が存在する。Lane らは入力ベクトルや学習パラメータについて微分可能な高次 CMAC (Higher-Order CMAC) を提案した [5]。宮本らは関数要素の線形和によって逆動力学モデルの学習を行っている [6]。これらのモデルでは、

$$\Phi(W, y) = M(W)h(y) \quad (29)$$

と書くことができる。\$h(y_d)\$ は関数要素から構成される \$p\$ 次元 (\$p > m\$) ベクトルである。\$m\$ 行 \$p\$ 列の行列 \$M(W)\$ の各成分は、例えば、次式のように定義される。

$$M_{ij} = W_{pi+j} \quad (30)$$

このような学習要素では、

$$B_w = \begin{bmatrix} h^T(y_d) & 0 \cdots 0 & \cdots & 0 \cdots 0 \\ 0 \cdots 0 & h^T(y_d) & \cdots & 0 \cdots 0 \\ \vdots & \vdots & & \vdots \\ 0 \cdots 0 & 0 \cdots 0 & \cdots & h^T(y_d) \end{bmatrix} \quad (31)$$

となり、

$$AB_w B_w^T A^+ = h^T(y_d)h(y_d)I_n \quad (32)$$

が成立する。\$I_n\$ は \$n\$ 行 \$n\$ 列の単位行列である。明らかに、式 (28) が成立し、対象システムの入力の次元が出力の次元よりも多い場合にも、目標指向性が存在することがわかる。

4. む す び

本論文では、対象システムの入力の次元と出力の次元が等しく、逆モデルの出力する近似解について対象のヤコビ行列がフルランクの場合、十分正確な逆モデルが学習された時点で、直接逆モデリングのオンライン

ン学習に目標指向性が存在することを示した。対象システムが1対1対応システムの場合、直接逆モデリングのオフライン学習によって、十分正確な逆モデルを学習することは可能であり、その時点において、オンライン学習を開始すれば、特定の目標について、目標指向性が保証される。

しかしながら、対象システムが多対1対応システムの場合、直接逆モデリングによる学習によって、十分に正確な逆モデルを学習できる保証はない。特殊な場合において、目標指向性が存在するとしても、一般のシステムの逆モデルの学習法としては、直接逆モデリングは不十分な手法である。一般のシステムの逆モデルをオンラインで学習する手法としては、順逆モデリング[1]や順逆モデリングの弱点を改良するためにフィードバック誤差学習則[6]を拡張した拡張フィードバック系を利用した逆モデル学習法[7]が望ましいと考える。

文 献

- [1] M.I. Jordan, "Supervised learning and systems with excess degrees of freedom," COINS Technical Report, 88-27, 1-41, 1988.
- [2] J.S. Albus, "A new approach to manipulator control: The cerebellar model articulation controller (CMAC)," Transactions of the ASME, J. Dynamic Systems, Measurement, and Control, 97, pp.220-227, 1975.
- [3] S.I. Colombano, M. Compton, and M. Baulat, "Goal directed model inversion," Proc. of International Joint Conference on Neural Networks (IJCNN'91 Singapore), pp.2422-2427, 1991.
- [4] D.E. Rumelhart, G.E. Hinton, and R.J. Williams, "Learning Internal Representation by Error Propagation," Parallel Distributed Processing, eds. D.E. Rumelhart and J.L. McClelland and the PDP Research Group, pp.318-326, MIT Press, 1986.
- [5] S.H. Lane, D.A. Handelman, and J.J. Gelfand, "Theory and development of higher-order CMAC neural networks," IEEE Control Systems, vol.14, no.2, pp.23-30, 1992.
- [6] H. Miyamoto, M. Kawato, T. Setoyama, and R. Suzuki, "Feedback-error-learning neural network for trajectory control of a robotic manipulator," Neural Networks, vol.1, pp.251-265, 1988.
- [7] 大山英明, 館 暉, "拡張フィードバック系による逆モデル学習," 信学論 (D-II), vol.J-78-D-II, no.2, pp.349-357, Feb. 1995.
(平成8年4月30日受付, 10月16日再受付)