ニューラルネットワークによるサーボ系の
ディジタル制御*
(オフライン学習の一手法)劉 明 輝^{*1}, 藤 堂 勇 雄^{*2}Digital Control of Servo Systems Using Neural Networks
(A Method of Off-Line Learning)

Ming-hui LIU and Isao TODO

When neural networks are applied to servo systems, the computation time of the control algorithms cannot be neglected, and the time period of on-line learning is expected to be shortened, considering the life span of the plant and the saving of electric energy. In this paper, we first proposed a design method of digital control systems using neural networks. In this method, the future information of the desired value is used as the input of the neural network and the computation time of the control algorithm is taken into account. Secondly, we presented an off-line learning method based on an approximate model of the servo system. The effectiveness of the proposed control systems and the off-line learning method is demonstrated by experiments and simulations on the control of a parallelogram link robot manipulator of 2 degrees of freedom.

Key Words: Automatic Control, Robotics, Control Device, Digital Control, Robot Manipulator, Neural Network, Learning

1. ま え が き

最近、ニューラルネットワーク（以下、NNと略す）をサーボ系の制御に応用する研究が盛んに行われている。このようなNNを用いて制御系を構成する場合、制御アルゴリズムとしては、(i)連続時間形式（アナログ式）のものと(ii)離散時間形式（ディジタル式）のものがある。

これらの制御アルゴリズムをマイクロプロセッサで実現するには、連続時間形式のアルゴリズムの場合には、これを離散時間形式のものに変換しなければならない。これまで、連続時間形式で表した制御アルゴリズムの提案⁽¹⁾⁽²⁾は多いが、最初から離散時間形式で表した制御アルゴリズムの研究は少ない⁽³⁾。

一方、このようなNNを用いる学習制御系では、制御量と目標値との誤差が許容範囲内に入るまで長時間の学習が必要である。ロボットなどサーボ系の制御に、NNを用いた制御方法を適用しオンライン学習を行うと、(i)学習の初期段階では誤差が大きいの

で、(ii)制御対象の寿命が短くなること、(iii)電力の消費量が多いことなどの点で問題がある。

本研究では、まず、制御アルゴリズムを最初から離散時間形式で表し、NNを用いたディジタル制御系の一構成法を提案した。この制御系は、目標値の予見情報を利用し、NNを計算する際の演算遅れを考慮して制御系を構成するものである。すなわち、(i)サンプリング周期 T を演算時間より長く選ぶ場合、および、(ii)演算時間を $2T$ の間にわたって配分する場合について、制御系の構成法を提案した。

ついで、制御対象の近似モデルによるオフライン学習の一手法を提案した。このオフライン学習の手法は、特性未知のサーボ系に試験信号を加えて近似モデルを作成し、これを制御対象の代わりに制御系に組込んでNNにオフライン学習させたのち、実機でオンライン学習を行うというものである。平行リンク機構をもつ2自由度ロボットマニプレータの軌道制御に本手法を適用し、その有効性を実験とシミュレーションによって確かめた。

2. フィードバック制御装置とNNの設計

提案した制御系の基本構成を図1に示す。以下、こ

* 平成2年7月14日 機械力学講演会において講演、原稿受付 平成2年8月3日。

^{*1} 正員、横浜国立大学大学院 (〒240 横浜市保土ヶ谷区常盤台156)。

^{*2} 正員、横浜国立大学工学部。

の制御系を構成1と呼ぶ。図1は、制御対象としてロボットマニプレータを想定し、その1個のリンクを制御する場合の構成である。 $G_c(z)$ は従来の線形フィードバック制御装置であり、本研究では、PD制御装置とする。NNはニューラルネットワークによる制御装置である。 $r(k)$, $y(k)$, $u(k)$, $v(k)$ は、それぞれ目標値、制御量、操作量、NNの出力であり、NNへの入力は、目標値とその予見値および制御量からなっている。 K はNNの出力を調整する定数であり、NNの学習はPD制御装置の出力 $m(k)$ によって駆動される。

本研究の設計は、二つの段階に分けられる。

第1段階では、PD制御装置 $G_c(z)$ を設計する。系の安定性と外乱に対するロバスト性は、このフィードバックループによって確保される。

第2段階では、NN制御装置を設計する。学習によってNN制御装置は制御対象の逆システムまたは近似逆システムに近づく。

2.1 フィードバック制御装置の設計 まず、図1のフィードバック制御装置を設計する。本研究では、 $G_c(z)$ を

$$G_c(z) = K_p + K_d(1 - z^{-1}) \quad (1)$$

とおく。

$G_c(z)$ は、制御系の安定性、ロバスト性、NNの学習速度や収束状況などに影響があるから、これらのことを考慮して、 K_p と K_d の値を決定する。

2.2 NNの設計 実験を行う際の演算時間を考慮して、少ない個数のニューロンでNNを構成する。図1のNNは、入力層4個、中間層6個、出力層1個のニューロンから構成されている。ただし、入力層(A層)の i 番目のニューロンの出力を y_i^A 、中間層(B層)の j 番目の入出力を x_j^B , y_j^B 、出力層(C層)の入出力を

x^C , v_o とし、 y_i^A と x_j^B とは重み w_{ij}^{AB} , y_j^B と x^C とは w_j により結合しており、同一層内のユニット間の結合や入力層から出力層までの直接結合はないとする。入力層は線形で、入力層のニューロンの出力は入力と同一である。中間層と出力層の入出力には次の関係があるものとする。

$$\left. \begin{aligned} x_j^B &= \sum_i w_{ij}^{AB} y_i^A, \quad y_j^B = f(x_j^B) \\ x^C &= \sum_j w_j y_j^B, \quad v_o = f(x^C) \end{aligned} \right\} \quad (2)$$

ただし

$$f(x) = \frac{2}{1 + e^{-x}} - 1 \quad (3)$$

w_j , w_{ij}^{AB} ($n=0, 1, 2, \dots$)の学習には、次の誤差逆伝ば法を用いる。すなわち、まず、

$$\left. \begin{aligned} \delta &= K f'(x^C) m(k) \\ w_j(n+1) &= w_j(n) + \eta \delta y_j^B \\ (j &= 1, \dots, 6) \end{aligned} \right\} \quad (4)$$

を用いて、 w_j を算出し、ついで、

$$\left. \begin{aligned} \delta_j &= f'(x_j^B) \delta w_j(n) \\ w_{ij}^{AB}(n+1) &= w_{ij}^{AB}(n) + \eta \delta_j y_i^A \\ (i &= 1, \dots, 4; j = 1, \dots, 6) \end{aligned} \right\} \quad (5)$$

を用いて w_{ij}^{AB} を求める。ただし、 η は学習率、 $f'(x) = df(x)/dx$ である。

3. 演算遅れを考慮した制御系の構成

従来の研究^{(1)~(3)}のほとんどは、NNの演算時間を問題とせず、それに関する計算が十分速く行えることを前提としている。しかしながら、実際には、制御装置の演算遅れが問題になる。そこで、本研究では、NNを含む制御装置の演算遅れを考察した制御系の二つの構成法をまず提案する。

図1 提案した制御系の基本構成(構成1)

図 2 異なるサンプリング周期をもつ制御系の構成 (構成 2)

3・1 構成 1 図 1 の構成 1 では, NN の学習 (すなわち, 重み更新) はサンプリング周期ごとに行うため, NN を含む制御則の計算はすべて 1 サンプル周期内に終わらなければならない。構成 1 では, 演算遅れの影響を考慮し, $r(k+1)$, $r(k+2)$, $r(k+3)$ をニューラルネットワークへの入力として用いた。

3・2 構成 2 演算時間が長いとき, 図 2 の構成 2 を提案する。構成 2 では, 最初のサンプリング周期で NN の出力を計算し, 次のサンプリング周期で重みの更新を行う。サンプラ S_1 は, $k=1, 3, 5, \dots$ のとき, S_2 は, $k=0, 2, 4, \dots$ のとき閉じるものとする。

3・3 演算時間の配分 図 3 (a), (b) は構成 1 と構成 2 の制御則を計算するときの時間配分である。ただし, t_1 は PD 制御装置 $G_c(z)$ を計算する時間, t_2 は NN の出力 $v_o(k)$ を計算する時間, t_3 は NN の重み更新に必要な時間であり, $t_d = t_1 + t_2 + t_3$ とする。本設計では, 演算時間 t_d の大きさによって, 構成 1 と構成 2 の 2 通りの制御系を構成することができる。

4. オフライン学習の一手法

図 1, 2 の制御系の構成は制御対象に関する知識をほとんど必要としない点では好都合であるが, ロボットなどのサーボ系の制御対象に長時間のオンライン学習をさせることは望ましくない。そこで本章では, この問題を解決するため, 図 4 に示すように, 制御対象の近似モデルによってオフライン学習を行うという一手法を提案する。

4・1 モデルの作成 0 形の制御対象の数式モデルをステップ応答から求める方法⁽⁴⁾ はよく知られているが, 本研究では, 図 5 (a) に示す幅 MT の試験信号を実在の制御対象に加えて速度応答 $x_v(k)$ を求め

(a) 構成 1 ($t_d < T$) (b) 構成 2 ($T < t_d < 2T$)

図 3 演算時間の配分

る。ただし, T はサンプリング周期である。制御対象を同図 (b) で近似すると, $g_k(k=1, 2, \dots)$ は

$$\left. \begin{aligned} g_k &= \frac{x_v(k) - x_v(k-1)}{A} & (k \leq M) \\ g_k &= \frac{x_v(k)}{A} - g_{k-m+1} - g_{k-m+2} - \dots - g_{k-1} & (M < k) \end{aligned} \right\} \dots\dots (6)$$

で与えられる。

図 5 (b) は同図 (c) の信号伝達線図で表現されるから, 制御対象は

$$\left. \begin{aligned} x(k+1) &= Px(k) + qu(k) \\ y(k) &= c^T x(k) \end{aligned} \right\} \dots\dots\dots (7)$$

で表される。ただし

$$P = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 \\ Tg_n & Tg_{n-1} & \cdots & Tg_2 & Tg_1 & 1 \end{bmatrix}$$

$$q = [0 \ 0 \ \cdots \ 0 \ 1 \ 0]^T$$

$$c^T = [Tg_n \ Tg_{n-1} \ \cdots \ Tg_1 \ 1]$$

.....(8)

4.2 オフライン学習 まず、モデル同定で求めた近似モデル式(7)を制御対象にして、制御系をオフラインで学習させたのち、制御対象を実機に切換えてオンライン学習を行う。

5. 実験

本論文で提案した構成法の有効性を確かめるため、

目標値に対するロボットマニプレータの手先の応答を実験によって求めた。使用したロボットマニプレータの略図を図6に示す。マニプレータは2個の直流モータ1, 2からなり、リンク1はモータ1、リンク2はモータ2により駆動される。両モータの回転角度を $y_1(k), y_2(k)$ とおく。マニプレータは減速機を用いているので、モータ1とモータ2の間の干渉を無視することができる。このため、制御系は二つの独立な1入力1出力系と考える。 r, u, v, y の添字1, 2はそれぞれリンク1, 2を表す。 $K=25, \eta=0.003$ とする。構成1, 構成2の制御系のパラメータを表1に示す。

5.1 オンライン学習 本実験では、制御アルゴリズムの演算時間は $t_1=0.55$ ms, $t_2=3.0$ ms, $t_3=2.3$ ms, $t_d=5.85$ msであった。このため、構成1では、 $T=10$ ms, 構成2では $T=5$ msに選んだ。構成1, 2の実験結果をそれぞれ図7, 8に示す。図7(c)と図8の縦軸は誤差の二乗平均 $\overline{e^2}$, 横軸は学習回数であり、

図4 オフライン学習による制御系(構成1)

(a) 試験信号

(b) 近似モデル

(c) 制御対象の信号伝達線図

図5 サーボ系の近似モデル

円の直径は 10 cm である。フルストップ誤差が大きい
うちは PD フィードバック制御が中心で、NN の出力
は 0 に近いが、学習するに従って誤差が小さくなり、
NN による制御が主制御となる。どちらの構成でもほ
ぼ 30 回の学習ののち、出力は目標値とほとんど同じ
になる。図 8 は図 7 (c) とほぼ一致している。しかし
ながら、構成 2 では PD フィードバック制御ループの
サンプリング周期は構成 1 より小さいため、構成 2 は
外乱に対してすぐれているといえる。

5・2 構成 1 のオフライン学習 図 9 (a), (b)

図 6 ロボットマニプレータの略図

(a) リンク 1 (50 回めの試行)

(b) リンク 2 (50 回めの試行)

(c) 学習過程

図 7 オンライン学習 (構成 1)

は、それぞれリンク 1、リンク 2 の試験信号とその応答である。

図 10 はオフライン学習の実験結果である。同図 (a), (b) は、それぞれリンク 1、2 の応答を示す。同図 (c), (d) はオフラインの学習過程を示す。

50 回のオフライン学習ののち、学習を行った以外
の入力を加えて応答を調べた。図 11 (a) は辺長 10 cm
の正方形、同図 (b) は直径 14 cm の円に対する実機での
1 回めの応答である。

図 10, 11 からオフライン学習の効果がよくわかる。

表 1 制御系のパラメータ

	Case 1	Case 2
T	0.01 s	0.005 s
K_P	10 V/rad	10 V/rad
K_D	100 V/rad	200 V/rad
M	10	20
n	20	40
A	3.0 V	3.0 V

図 8 オンライン学習 (構成 2)

(a) リンク 1

(b) リンク 2

図 9 試験信号と応答

(a) リンク 1 (50 回目の試行)

(b) リンク 2 (50 回目の試行)

(c) リンク 1 (d) リンク 2

図 10 オフライン学習 (構成 1)

(a) 正方形 (b) 円

図 11 オフライン学習後の 1 回目の試行 (構成 1)

(a) リンク 1

(b) リンク 2

図 12 オフライン学習 (構成 2)

5.3 構成 2 のオフライン学習 本オフライン学習法は構成 2 にも拡張することができる。その実験結果を図 12 に示す。以上の実験結果から、提案したオフライン学習法が有効であることがわかる。

6. シミュレーション

実験と比べるため、ロボットマニピュレータをマイクロコンピュータでシミュレートする。ロボットは連続時間系モデルであるため、図 13 のようにして離散時間系モデルを求める。このとき、ロボットのパルス伝達関数 $G(z)$ は

$$G(z) = \frac{qz(\beta z^{-1} + \gamma z^{-2})}{1 - (p+1)z^{-1} + pz^{-2}} \quad \dots\dots\dots (9)$$

で表される。ただし

図 13 ロボットモデルの離散化

(a) 構成 1

(b) 構成 2

図 14 オンライン学習 (シミュレーション)

$$\left. \begin{aligned} p &= e^{-aT} \\ q_1 &= \frac{a}{b} \left(T - \frac{1-p}{a} \right) \\ q_2 &= \frac{b(1-p)}{a} \\ \beta &= \frac{q_1}{q_2} \\ \gamma &= \frac{1-p}{a} - \beta p \end{aligned} \right\} \dots\dots\dots (10)$$

周波数応答法によって実機の a, b の値を推定した。その結果、リンク 1 については、 $a=26.59, b=10.0$ 、リンク 2 については、 $a=28.98, b=10.9$ が得られた。

式(9)のモデルを使った構成 1 と構成 2 のシミュレーション結果を図 14 に示す。図 14(a), (b) は、それぞれ図 7(c), 図 8 の実験とほぼ一致していることがわかる。

7. ま と め

本論文の主要な結果は次のとおりである。

(1) NN を用いたデジタル制御系の構成を提案し、その効果を実験とシミュレーションによって実証した。

(2) NN を用いてサーボ系をデジタル制御する場合のオフライン学習の一手法を提案し、その効果

を実験によって実証した。

(3) 構成 2 では、PD フィードバック制御系のサンプリング周期を小さくできるという特長があり、この構成はフィードバック制御装置が有効に働いているときは有利である。

(4) 本手法を用いることによって、実機でのオンライン学習回数を減らすことが可能となった。

終わりに、本研究は、(財)メカトロニクス技術高度化財団の研究助成を受けて進められたことを付記し、謝意を表す。また、実験を行うに当たって便宜を図ってくださったトキコ(株)研究所・内海典之主任研究員に感謝する。

文 献

- (1) 川人・宇野・鈴木, 随意運動における適応と学習 II, 日本ロボット学会誌, 6-3(1988), 50.
- (2) 瀬戸・川人・鈴木, 多層神経回路網内に学習される逆ダイナミクスモデルによるマニプレータの制御, 信学技報, MBE 87(1988), 249.
- (3) 山田・藪田, ニューラルネットを用いたダイレクトコントローラ, 計測自動制御学会学術講演会予稿集, (1989-7), 391.
- (4) Takahashi, Y., Tomizuka, M. and Auslander, D. M., Simple Discrete Control of Industrial Processes, *Trans. ASME, J. Dyn. Syst. Meas. Control*, 97-4(1975), 354.