Sensors & Transducers, Vol. 21, Special Issue, May 2013, pp. 142-146

Sensors & Transducers

http://www.sensorsportal.com

Convergence Analysis of Multi-innovation Learning Algorithm Based on PID Neural Network

[1] Gang Ren, [1,2] Pinle Qin, [2] Minmin Sun, [1] Yan Lin
[1] Ship CAD Engineer Center, Dalian University of Technology, Liaoling, China
[2] Department of Computer Science, North University of China, Shanxi, China
[1] E-mail: pinleqin@126.com

Received: 26 March 2013 /Accepted: 14 May 2013 /Published: 30 May 2013

Abstract: In order to improve the identification accuracy of dynamic system, multi-innovation learning algorithm based on PID neural networks is presented, which can improve the online identification performance of the networks. The multi-innovation gradient type algorithms use the current data and the past data that make it more effective than the BP algorithm in view of accuracy and convergence rate. Simulation results showed that the proposed algorithm is effect. *Copyright © 2013 IFSA.*

Keywords: Multi-innovation, PID neural networks, System identification, Nonlinear system.

1. Introduction

As we have known, the model-based control techniques are usually implemented under the assumption of good understanding of process dynamics and they are the dependence on mathematical model of controlled plant. These techniques, however, can not provide satisfactory results when applied to poorly modeled processes. Therefore, how to design adaptive control system only based on information from the I/O data of the plant is great significance both in theoretic and application. Now, one successful method of the model-free control approaches is Proportional Integral Derivative neural network (PIDNN) control [1, 2].

PIDNN is a new kind of networks. It utilizes the advantages of both PID control and neural structure. But the method has some shortcomings, such as slow convergence, easy to fall into local minimum and forget the samples [3, 4].

In order to overcome these shortcomings, this paper improves the convergence by using multi innovation theory [5, 6] and proofs the fast convergence of the improvement method by supermartigale theory.

2. PID Neural Network Model and Multi-innovation Identification Algorithm

2.1. PID Neural Network Model

Consider a time-invariant stochastic system described by a linear regression model [7, 8]:

$$y(t)=\phi^T(t)\theta+v(t), \tag{1}$$

where $\theta\in R^n$ is the parameter vector of system, $y(t)\in R^1$ is the system output, $u(t)\in R^1$ is the system input, $v(t)\in R^1$ is a stochastic noise with zero mean, $\phi(t)\in R^n$ is the information vector consisting

Article number P_SI_356

of the system observation (input-output) data, the superscript T denotes the transpose.

As it is shown in Fig. 1, PID neural network has a simple feed forward neural network which consists of 2-3-1 structure, so it has three layers [7-9]. Let $W=(w_{ij})_{2\times 3}$ be the weight matrix between the input layer and the hidden layer and $V=(v_1,v_2,v_3)$ be the weight matrix between the hidden layer and the output layer. u_i, u'_j and u'' are the inputs of the input layer, the hidden layer and the output layer. x_i, x'_j and x'' are the outputs of them. The actual output of the PID neural network is d.

We can use Fig. 1 to approximate the process described by expressing (1).

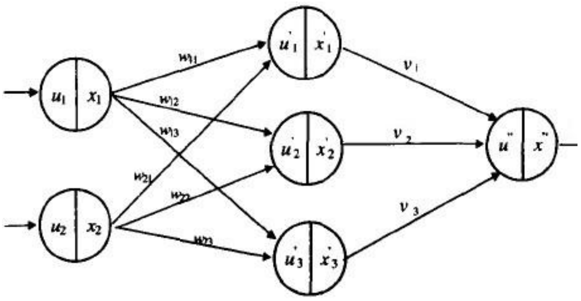

Fig. 1. Structure of PIDNN.

2.2. Multi-innovation Identification Algorithm

The following single innovation identification algorithm can be used to estimate the parameter θ of system (1) [10].

$$\hat{\theta}(t)=\hat{\theta}(t-1)+\frac{\phi(t)}{r(t)}e(t) \tag{2}$$

$$e(t)=y(t)-\phi^T(t)\hat{\theta}(t-1) \tag{3}$$

$$r(t-1)=r(t-1)+\|\phi(t)\|^2,\quad r(0)=1, \tag{4}$$

where $e(t)$ is a single innovation.

In multi-innovation identification algorithm [5], the single innovation $e(t)\in R^1$ is extended to multi-innovation vector $E(p,t-1)\in R^p$, the information vector $\phi(t)\in R^n$ is also extended to $\Phi(p,t)\in R^{n\times p}$ for scalar system. From here, the algorithm can be expressed as

$$\hat{\theta}(t)=\hat{\theta}(t-1)+\frac{\Phi(p,t)}{r(t)}E(p,t) \tag{5}$$

$$E(p,t-1)=[e(t-1),e(t-1),\cdots,e(t-p)]^T\in R^p \tag{6}$$

$$\Phi(p,t)=[\phi(t),\phi(t-1),\cdots,\phi(t-p+1)]\in R^{n\times p} \tag{7}$$

where $p\geq 1$ represents the innovation length. As $p=1$, the multi-innovation identification algorithm reduces to the single innovation identification algorithm.

3. Multi-innovation Learning Algorithm based on PID Neural Network

Considering p groups of input-output data from time $t-p+1$ to time t, the output vectors of the ith input node is

$$X_i(p,t)=[x_i(t),x_i(t-1),\cdots,x_i(t-p+1)],\quad i=1,2$$

The output vector of the output node is

$$D(p,t)=[d(t),d(t-1),\cdots d(t-p+1)]^T$$

The expected vector of PID neural network is

$$Y(p,t)=[y(t),y(t-1),\cdots,y(t-p+1)]^T$$

Define a cost function [10]:

$$J(p,t)=\frac{1}{2}\|Y(p,t)-D(p,t)\|^2 \tag{8}$$

Thus, we proceed to refine the weights by the training iteration as follows:

$$v_j(t+1)=v_j(t)-\eta\frac{\partial J(p,t)}{\partial v_j}$$

$$w_{ij}(t+1)=w_{ij}(t)-\eta\frac{\partial J(p,t)}{\partial w_{ij}} \tag{9}$$

where $\eta>0$ is the learning rate.

$$\begin{aligned}\frac{\partial J(p,t)}{\partial v_j}&=\frac{\partial X''(p,t)}{\partial v_j}\frac{\partial J(p,t)}{\partial X''(p,t)}\\&=X'_j(p,t)[\frac{\partial J(p,t)}{\partial x''(t)},\cdots,\frac{\partial J(p,t)}{\partial x''(t-p+1)}]^T\\&=X'_j(p,t)[\frac{\partial J(p,t)}{\partial d(t)},\cdots,\frac{\partial J(p,t)}{\partial d(t-p+1)}]^T\\&=X'_j(p,t)[-e(t),\cdots,-e(t-p+1)]^T\\&=-X'_j(p,t)E(p,t)\end{aligned}$$

by defining the vector $X'_j(p,t)$ as

$$X_j'(p,t)=[x_j'(t),x_j'(t-1),\cdots,x_j'(t-p+1)]$$

and

$$\begin{aligned}\frac{\partial J(p,t)}{\partial w_{ij}}&=\frac{\partial U_j'(p,t)}{\partial w_{ij}}\frac{\partial J(p,t)}{\partial U_j'(p,t)}\\&=X_i(p,t)[\frac{\partial J(p,t)}{\partial u_j'(t)},\cdots,\frac{\partial J(p,t)}{\partial u_j'(t-p+1)}]^T\\&=X_i(p,t)[-e_1(t),\cdots,-e_1(t-p+1)]^T\\&=-X_i(p,t)E_1(p,t)\end{aligned}$$

the vector $X_j'(p,t)$ and $e_1(t)$ may be expressed as

$$X_i(p,t)=[x_i(t),x_i(t\text{-}1),\cdots,x_i(t-p+1)]$$

$$\begin{aligned}e_1(t)&=-\frac{\partial J(p,t)}{\partial d(t)}\frac{\partial d(t)}{\partial x_j'(t)}\frac{\partial x_j'(t)}{\partial u_j'(t)}\\&=e(t)v_j(t)sign\frac{x_j'(t)-x_j'(t-1)}{u_j'(t)-u_j'(t-1)}\end{aligned}$$

So we have

$$v_j(t+1)=v_j(t)+\eta X_j'(p,t)E(p,t)$$

$$w_{ij}(t+1)=w_{ij}(t)+\eta X_i(p,t)E_1(p,t) \qquad (10)$$

4. Convergence Analysis of multi-Innovation Learning Algorithm

For the identification of nonlinear dynamical systems, PID neural network can be treated as a dynamical system described by (1). Its adjustable parameters $\theta=(w_{ij},v_j)$ which is made of $W=(w_{ij})_{2\times3}$ and $V=(v_1,v_2,v_3)$ can be written as (5).

Reference [11, 12], convergence analysis of the proposed algorithm be given as fellows.

Theorem 1. For the system in (1), assume that a stochastic white noise with zero mean $v(t)$ is uncorrelated with $u(t)$

(A1) $E[v^2(t)]\le\sigma_v^2<\infty$.

if there exist constants $0<\alpha\le\beta<\infty$ and the innovation length $p\ge n$ such that the following persistent excitation condition holds,

(A2) $$\alpha I\le\frac{1}{p}\sum_{i=1}^{p}\phi(t-i+1)\phi^T(t-i+1)\le\beta I.$$

Then the parameter estimation $\hat\theta(t)$ given by Eq.(5) is uniform bounded.

Proof. Define the estimation error $\tilde\theta(t)=\theta(t)-\theta$, assume θ is real value. Subtracting θ from the two sides of Eq.(5), we have

$$\begin{aligned}\theta(t)-\theta&=\theta(t-1)-\theta+\frac{\Phi(p,t)}{r(t)}[\Phi^T(p,t)\theta+\\&V(p,t)-\Phi^T(p,t)\theta(t-1)]\\&=\theta(t-1)-\theta+\frac{\Phi(p,t)}{r(t)}[V(p,t)-\\&\Phi^T(p,t)(\theta(t-1)-\theta)]\end{aligned} \qquad (11)$$

Substituting $\tilde\theta(t)$

$$\begin{aligned}\tilde\theta(t)&=\tilde\theta(t-1)+\frac{\Phi(p,t)}{r(t)}[-\Phi^T(p,t)\tilde\theta(t-1)+V(p,t)]\\&=\left[I-\frac{\Phi(p,t)\Phi^T(p,t)}{r(t)}\right]\tilde\theta(t-1)+\frac{\Phi(p,t)V(p,t)}{r(t)}\end{aligned} \qquad (12)$$

By using A2, taking the trace gives

$$p\alpha\le\frac{1}{p}\sum_{i=1}^{p}\|\phi(t-i+1)\|^2\le p\beta$$

$$np\alpha\le\sum_{i=1}^{p}\|\phi(t-i+1)\|^2\le np\beta$$

Hence, we have

$$np\alpha\le\|\phi(t)\|^2\le np\beta$$

According to the definition of $r(t)$ in (4), we have

$$\begin{aligned}r(t)&=\sum_{j=1}^{t}\|\phi(j)\|^2+r(0)\\&\le\sum_{j=0}^{\lfloor(t-1)/p\rfloor}\sum_{i=1}^{p}\|\phi(jp+i)\|^2+r(0)\\&\le\sum_{j=0}^{\lfloor(t-1)/p\rfloor}np\beta+1\\&\le(\lfloor(t-1)/p\rfloor+1)np\beta+1\\&\le(t+p-1)n\beta+1\end{aligned}$$

$$\begin{aligned}r(t)&=\sum_{j=1}^{t}\|\phi(j)\|^2+r(0)\\&\ge\sum_{j=0}^{\lfloor t/p\rfloor-1}\sum_{i=1}^{p}\|\phi(jp+i)\|^2+r(0)\\&\ge\sum_{j=0}^{\lfloor t/p\rfloor-1}np\alpha+r(0)\\&\ge(\lfloor t/p\rfloor-1+1)np\alpha+1\\&\ge tn\alpha+1\end{aligned}$$

Then, the following inequality has been obtained

$$tn\alpha + 1 \le r(t) \le (t+p-1)n\beta + 1$$

$$I - \frac{\Phi(p,t)\Phi^T(p,t)}{r(t)} \le \left[1 - \frac{p\alpha}{(t+p-1)n\beta+1}\right] I$$

Here, using assume A1, we have

$$\begin{aligned} E(\|\Phi(p,t)V(p,t)\|) &\le \sqrt{E(\|\Phi(p,t)V(p,t)\|^2)} \\ &\le \sqrt{p\beta E(\|V(p,t)\|^2)} \\ &\le \sqrt{p\beta p\sigma_v^2} \\ &\le p\sigma_v\sqrt{\beta} \end{aligned} \tag{13}$$

Taking the expectation and norm of (12) and using (13) give

$$\begin{aligned} \|\tilde{\theta}(t)\| &= \left\| \left[I - \frac{\Phi(p,t)\Phi^T(p,t)}{r(t)} \right] \tilde{\theta}(t-1) + \frac{\Phi(p,t)V(p,t)}{r(t)} \right\| \\ &\le \left\| \left[I - \frac{\Phi(p,t)\Phi^T(p,t)}{r(t)} \right] \tilde{\theta}(t-1) \right\| + \left\| \frac{\Phi(p,t)V(p,t)}{r(t)} \right\| \\ &\le \left\| 1 - \frac{p\alpha}{(t+p-1)n\beta+1} \right\| \|\tilde{\theta}(t-1)\| + \frac{\|\Phi(p,t)V(p,t)\|}{r(t)} \\ &\le \left\| 1 - \frac{p\alpha}{(t+p-1)n\beta+1} \right\| \|\tilde{\theta}(t-1)\| + \frac{\|\Phi(p,t)V(p,t)\|}{r(t)} \\ &\le \left\| 1 - \frac{p\alpha}{(t+p-1)n\beta+1} \right\| \|\tilde{\theta}(t-1)\| + \frac{\|\Phi(p,t)V(p,t)\|}{tn\alpha+1} \end{aligned}$$

$$\begin{aligned} E\left(\|\tilde{\theta}(t)\|\right) &\le \left\| 1 - \frac{p\alpha}{(t+p-1)n\beta+1} \right\| E\left(\|\tilde{\theta}(t-1)\|\right) + \\ &\quad \frac{1}{tn\alpha+1} E\left(\|\Phi(p,t)V(p,t)\|\right) \\ &\le \left\| 1 - \frac{p\alpha}{(t+p-1)n\beta+1} \right\| E\left(\|\tilde{\theta}(t-1)\|\right) + \frac{p\sigma_v\sqrt{\beta}}{tn\alpha+1} \end{aligned}$$

If we select suitable parameter α and β, it follows that

$$0 < 1 - \frac{p\alpha}{(t+p-1)n\beta+1} \le d < 1$$

where d is a constant. Thus,

$$E\left(\|\tilde{\theta}(t)\|\right) \le dE\left(\|\tilde{\theta}(t-1)\|\right) + M \tag{14}$$

where M is a constant. Repeating using the formula (14)

$$\begin{aligned} E\left(\|\tilde{\theta}(t)\|\right) &\le dE\left(\|\tilde{\theta}(t-1)\|\right) + M \\ &\le d^{t-1}E\left(\|\tilde{\theta}(1)\|\right) + \frac{1-d^{t-1}}{1-d}M \end{aligned}$$

This completes the proof of Theorem 1.

5. Examples

To illustrate the effectiveness of the proposed algorithm, this paper compares its accuracy with the traditionaly BP algorithm with the following example.

Example 1. Considering the nonlinear dynamical systems described by the following function.

$$y(t) = 0.4y(t-1) + 0.54y(t-2) + f[u(k-1)]$$

where $y(t)$ is the output system, $u(t)$ is the input system,

$$f(u) = u^3 + u^2 - 2.5u$$

The test signal used for the example is

$$u(t) = 0.2\sin\frac{2\pi t}{25} + 0.3\sin\frac{\pi t}{75}$$

The initial weight matrix in output layer $W(0) = [1, 0.1, 1; -1, -0.1, -1]$, the initial weight matrix in hidden layer $V(0) = (0.1, 0.1, 0.1)$, the innovation length $p = 2$, the learning rate $\eta = 0.175$. The system responses are shown in Fig. 2 and Fig. 3.

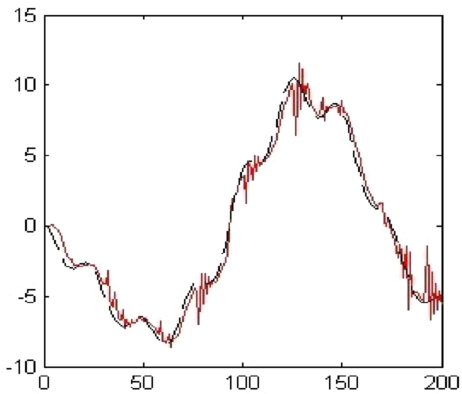

Fig. 2. Identification effect of the unmodified algorithm.

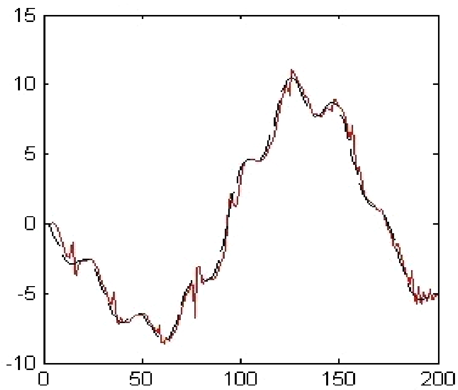

Fig. 3. Identification effect of the proposed algorithm.

Example 2. Considering the nonlinear dynamical systems described by the following function.

$$y(t+1) = f[y(t)] + g[u(k)]$$

where $f[y(t)] = \dfrac{5y(t)}{2.5 + y^2(t)}$, $g[u(t)] = u^3(t)$

The test signal used for the example is

$$u(t) = 0.6\sin\frac{2\pi t}{50} + 0.4\sin\frac{2\pi t}{75}$$

Here, the initial weight matrices of PID neural network be the same as the above example, the innovation length $p = 3$, the learning rate $\eta = 0.015$. The system responses are shown in Fig. 4 and Fig. 5.

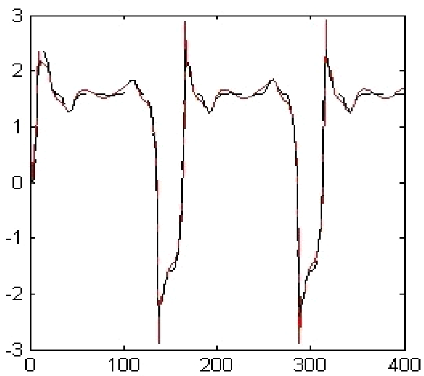

Fig. 4. Identification effect of the unmodified algorithm.

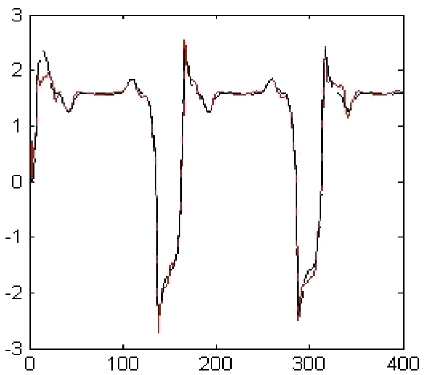

Fig. 5. Identification effect of the proposed algorithm.

6. Conclusions

This paper proposed a multi-innovation learning algorithm based on PID neural network and analyses the convergence of the algorithm. From the simulation example, it is shown that the proposed algorithm has fast convergence rate and good tracking ability.

References

[1]. Huailin Shu, Youguo Pi, PID neural networks for time-delay systems, *Computers and Chemical Engineering*, Vol. 5, 2000, pp. 859-862.

[2]. Shu Huai-Lin. Identification of Nonliner Systems Based on PID Neural Networks, *Acta Automatica Sinica*, Vol. 28, Issue 3, 2002, pp. 301-305.

[3]. Shou-Ju Li, Ying-Xi Liu, An Improved Approach to Nonlinear Dynamical System Identification using PID Neural Networks, *International Journal of Nonlinear Sciences and Numerical Simulation*, Vol. 7, Issue 2, 2006, pp. 177-182.

[4]. Xu Shaoyun, BP learning algorithm of online mode and batch mode, in *Proceedings of Chinese Intelligent Automation Conference Papers*, China Beijing, 15-17 May 1996, pp. 501-505.

[5]. Feng Ding, Tongwen Chen, Performance analysis of multi-innovation gradient type identification methods, *Automatica*, Vol. 43, Issue 1, 2007, pp. 1-17.

[6]. Liu Ying-Yu, Shen Dong-Ri, Chen Yi-Jun, Li Rong, Multi-innovation stochastic gradient identification algorithm based on feedforward neural networks, *Journal of Harbin University of Commerce (Natural Sciences Edition)*, Vol. 22, Issue 2, 2006. pp. 78-86.

[7]. Shu Huailin, On the Ability the PID Neural Networks in Nonlinear System Indentification, *Automation & Information Engineering*, Vol. 4, 2001, pp. 14-20.

[8]. Huailin Shu, Youguo Pi, Decoupled Temperature Control System Based on PID Neural Network, in *Proceeding of the ACSE' 05 Conference*, China Shanghai 14-16 August 2005, pp. 19-21.

[9]. Fardadi M., Selk Ghafari A., Hannani S. K., PID neural net-work control of SUT building energy management system, in *Proceedings of the IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, China Changsha 24-26 May 2005, pp. 682-686.

[10]. Qin Pinle, Lin Yan, Chen Ming, Improvement of Tracking Performance in Model-Free Adaptive Controller Based on Multi-innovation and Particle Swarm Optimization, *International Journal of Innovative Computing, Information and Control*, Vol. 5, Issue 5, 2009, pp. 1367-1377.

[11]. Lei Guo, Time-Varying Stochastic Systems-Stability, Estimation and Control, *Changchun: Jinlin Science and Technology Press*, 1993.

[12]. Yanjun Liu, Li Yu, Feng Ding, Multi-innovation extended stochastic gradient algorithm and its performance analysis, *Circuits Syst. Signal Processing*, Vol. 29, Issue 4, 2010, pp. 649-667.