

文章编号: 1001-0920(2005)10-1197-04

## PD 型 Elnan 网络及在动态系统辨识中的应用研究

漆为民, 程远楚, 姬巧玲, 蔡维由

(武汉大学 动力与机械学院, 武汉 430072)

摘 要: 首先介绍普通改进 Elnan 动态递归神经网络的结构, 重点讨论一种具有 PD 特性的 Elnan 神经网络及其学习算法, 并将改进 Elnan 网络和 PD 型 Elnan 网络分别用于动态系统的辨识。无论是理想的数学模型还是实际工业模型, 计算机仿真结果均证明, 将 PD 型网络用于动态系统辨识具有更好的逼近效果。

关键词: 动态递归神经网络; PD 型 Elnan 网络; 动态系统辨识

中图分类号: TP273

文献标识码: A

## PD Elnan Neural Network and Its Application to Dynamical System Identification

Q I W e i m i n , C H E N G Y u a n - c h u , J I Q i a o - l i n g , C A I W e i - y o u

(College of Power and Mechanical Engineering, Wuhan University, Wuhan 430072, China Correspondent: Q I W e i m i n , E m a i l : q w m - q w m @ 1 6 3 . c o m )

**Abstract** A modified Elnan dynamic recurrent neural network is introduced first. A PD Elnan neural network and its learning algorithm are discussed in detail. Elnan network and PD Elnan network are used to identify the dynamical system, respectively. Simulation results based on ideal mathematical model and industrial model show that the PD Elnan network is prior to the modified Elnan network in identification of nonlinear dynamic systems.

**Key words:** Dynamical recurrent network; PD Elnan neural network; Dynamical system identification

### 1 引 言

Elnan 神经网络是一种局部递归神经网络<sup>[1]</sup>, 它包含前馈和反馈两部分。前馈部分由输入层、隐层、输出层构成, 其连接权可学习修正; 反馈部分由一组“结构”单元构成, 用来记忆隐层单元前一刻的输出值, 其连接权是固定的。由于 Elnan 网络的动态特性仅由内部的连接提供, 无需使用状态作为输入或训练信号, 故其比静态前馈网络优越, 非常适合于动态系统辨识。然而, 基本 Elnan 采用标准 BP 学习算法训练时, 仅能辨识一阶线性动态系统。为了有效地辨识高阶系统, 人们提出了各种改进的 Elnan 网络, 理论分析证明它与基本 Elnan 网络相比具有更好的动态性能<sup>[2,3]</sup>。为了进一步改善 Elnan 网络的动态特性, 文献[4]提出了一种具有 PD 特性的改进 Elnan 神经网络。

本文在以上研究的基础上, 重点讨论其学习算法, 并将改进型和 PD 型 Elnan 网络分别用于动态系统的辨识。由仿真结果可以看出, PD 型 Elnan 网络具有更好的动态辨识性能。

### 2 改进 Elnan 动态递归网络

一种改进 Elnan 网络<sup>[3]</sup>如图 1 所示。其中:  $m, n, r$  分别表示网络的输入层、输出层、隐层的个数;  $w_1, w_2, w_3, w_4$  分别为结构单元到隐层、输入层到隐层、隐层到输出层、结构层到输出层的连接权矩阵;  $f(\bullet)$  和  $g(\bullet)$  分别为输出单元和隐层单元的激发函数所组成的非线性向量函数。

从图 1 可以看出, 改进 Elnan 网络将结构层的输出分别作为隐层和输出层的一个输入。令网络的输入  $u(k) \in R^m$ , 输出  $y(k) \in R^n$ , 隐层  $x(k) \in R^r$ , 则有

收稿日期: 2004-10-09; 修回日期: 2005-03-21

作者简介: 漆为民(1975-), 男, 武汉人, 博士生, 从事系统辨识与自适应控制、智能理论与智能控制的研究;  
程远楚(1962-), 男, 湖北仙桃人, 教授, 从事动力设备过渡过程与控制的研究

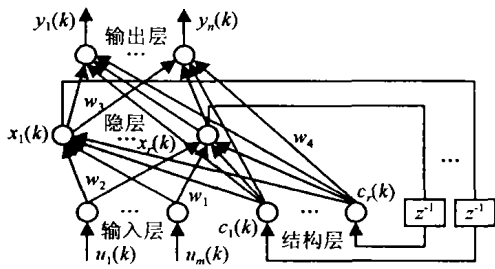


图1 一种改进的Elman网络结构图

$$x_j(k) = f \left[ \sum_{i=1}^m w_{2,i,j} u_i(k) + \sum_{i=1}^r w_{1,i,j} c_i(k) \right], \quad (1)$$

$$c_i(k) = x_i(k-1), \quad (2)$$

$$y_j(k) = g \left[ \sum_{i=1}^r (w_{3,i,j} x_i(k) + w_{4,i,j} c_i(k)) \right]. \quad (3)$$

研究证明,改进Elman网络具有比例积分特性,随着权系数的调整,其比例系数和积分系数不断发生变化,可以有效地辨识高于一阶的动态系统<sup>[2,3]</sup>.

### 3 PD型Elman动态递归网络

改进Elman网络具有比例积分特性,将其用于非线性系统具有较好的逼近性能.然而工业上大量使用PD型控制器,构造一种具有PD特性的人工神经网络,使之能更好地应用于工业实践已成为研究热点.参照Elman网络的基本思想,文献[4]提出了一种局部递归神经网络,如图2所示.

图2所示神经网络与前述改进Elman网络的不同之处在于增加了一个记忆层.记忆层从结构层获取信息,其输出分别作用于隐层和输出层.设结构层的自反馈系数为 $\alpha$ ,记忆层的自反馈系数为 $\beta$ ,结构层、输入层和记忆层至隐层的连接权矩阵分别为 $w_1, w_2$ 和 $w_6$ ;隐层、结构层和记忆层至输出层的连接权矩阵分别为 $w_3, w_4, w_5$ ;输入层、隐层、结构层、记忆层、输入层的状态分别为 $u, x, c, b, y$ .则其各层的非线性状态空间表达式为

$$x_j(k) = f \left[ \sum_{i=1}^m w_{2,i,j} u_i(k) + \sum_{i=1}^r w_{1,i,j} c_i(k) + \sum_{i=1}^r w_{6,i,j} b_i(k) \right]$$

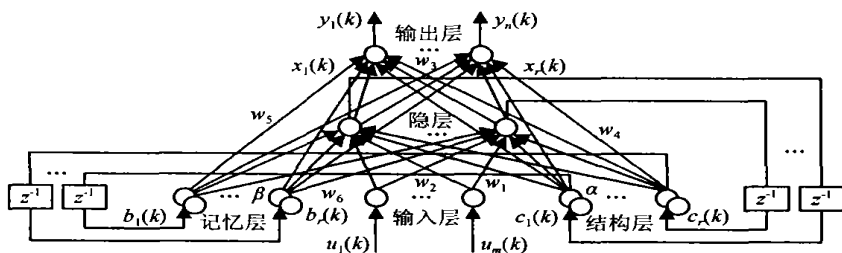


图2 PD型动态Elman神经网络

$$\left. \begin{aligned} & w_{1,i,j} b_i(k) \right\}, \quad (4) \end{aligned}$$

$$c_i(k) = x_i(k-1) + \alpha \times c_i(k-1), \quad (5)$$

$$b_i(k) = c_i(k-1) + \beta \times b_i(k-1), \quad (6)$$

$$y_j(k) = g \left[ \sum_{i=1}^r (w_{3,i,j} x_i(k) + w_{4,i,j} c_i(k) + w_{5,i,j} b_i(k)) \right]. \quad (7)$$

从系统的观点看,新的动态神经网络由两种网络构成,一种是前馈层,由输入层、隐层、结构层、记忆层和输出层组成;另一种是反馈层,由隐层、结构层和记忆层组成.前馈神经网络只利用了系统的静态信息,而新型神经网络在利用系统静态信息的同时,由于内部的存储能力,还能利用系统的动态信息;特别是通过两层记忆和双环反馈,更进一步增强了该网络的动态性能.文献[4]指出,新型神经网络具有对输入信号的比例、积分和微分作用,用于对系统进行辨识时必将加快收敛速度.

### 4 PD型Elman网络学习算法

从式(5)和(6)可知,结构层的信息与隐层节点前一时步的信息有关,而记忆层的信息又与结构层前一时步的信息相关,可以认为结构层 $c(k)$ 和记忆层 $b(k)$ 是一个动态递推过程.由此给出下面的反向传播学习算法.

设网络的理想输出为 $y_d$ ,网络的实际输出为 $y$ .考虑如下总体误差目标函数:

$$E(k) = \frac{1}{2} \sum_{p=1}^k \sum_{i=1}^n (y_{d_i}(p) - y_i(p))^2 = \sum_{p=1}^k E_p(k), \quad (8)$$

则网络的权值修正公式为

$$w_{1,j,i}(k+1) = w_{1,j,i}(k) + \eta \Delta w_{1,j,i}(k), \quad (9)$$

$$w_{2,j,q}(k+1) = w_{2,j,q}(k) + \eta \Delta w_{2,j,q}(k), \quad (10)$$

$$w_{3,i,j}(k+1) = w_{3,i,j}(k) + \eta \Delta w_{3,i,j}(k), \quad (11)$$

$$w_{4,i,j}(k+1) = w_{4,i,j}(k) + \eta \Delta w_{4,i,j}(k), \quad (12)$$

$$w_{5,i,j}(k+1) = w_{5,i,j}(k) + \eta \Delta w_{5,i,j}(k), \quad (13)$$

$$w_{6,j,i}(k+1) = w_{6,j,i}(k) + \eta \Delta w_{6,j,i}(k). \quad (14)$$

由于 $k-1$ 时刻的误差 $E(k-1)$ 反映在 $k$ 时刻的权值中,根据动态反向传播的思想, $k+1$ 时刻的

权值可用  $k$  时刻的误差  $E(k)$  进行修正, 即

$$\Delta w_{i,j} = - \frac{\partial E(k)}{\partial w_{i,j}} = \frac{\partial E_p(k)}{\partial w_{i,j}}, \quad (15)$$

则有

$$\Delta w_{3,i,j}(k) = - \frac{\partial E_p(k)}{\partial w_{3,i,j}(k)} = \delta_i(k) x_j(k) g_i'(\bullet), \quad (16)$$

式中  $\delta_i(k) = y_{di}(k) - y_i(k)$ . 同时有

$$\Delta w_{4,i,j}(k) = - \frac{\partial E_p(k)}{\partial w_{4,i,j}(k)} = \delta_i(k) c_j(k) g_i'(\bullet), \quad (17)$$

$$\Delta w_{5,i,j}(k) = - \frac{\partial E_p(k)}{\partial w_{5,i,j}(k)} = \delta_i(k) b_j(k) g_i'(\bullet). \quad (18)$$

由网络拓扑结构, 可以认为结构层  $c$  和记忆层  $b$  是网络的另两个隐层, 输出层  $k$  时步的误差也必然通过  $w_4$  和  $w_5$  反向传播回隐层去更新  $w_1, w_6$  和  $w_2$ .

对结构层到隐层的连接权  $w_1$  有

$$\Delta w_{1,j,l}(k) = \sum_{i=1}^n \delta_i(k) w_{3,i,j}(k) \frac{\partial \alpha_i}{\partial w_{1,j,l}} + \sum_{i=1}^n \delta_i(k) w_{4,i,j}(k) \frac{\partial \alpha_i}{\partial w_{1,j,l}} + \sum_{i=1}^n \delta_i(k) w_{5,i,j}(k) \frac{\partial \beta_i}{\partial w_{1,j,l}}. \quad (19)$$

式中

$$\frac{\partial \alpha_i(k)}{\partial w_{1,j,l}(k)} = f_j(\bullet) x_{li}(k-1) + \alpha \frac{\partial \alpha_i(k-1)}{\partial w_{1,j,l}(k-1)}, \quad (20)$$

$$\frac{\partial \alpha_i(k)}{\partial w_{1,j,l}(k)} = \frac{\partial \alpha_i(k-1)}{\partial w_{1,j,l}(k-1)}, \quad (21)$$

$$\frac{\partial \beta_i(k)}{\partial w_{1,j,l}(k)} = \frac{\partial \beta_i(k-2)}{\partial w_{1,j,l}(k-2)}. \quad (22)$$

类似地, 对记忆层到隐层的连接权  $w_6$  有

$$\Delta w_{6,j,l}(k) = \sum_{i=1}^n \delta_i(k) w_{3,i,j}(k) \frac{\partial \alpha_i}{\partial w_{6,j,l}} + \sum_{i=1}^n \delta_i(k) w_{4,i,j}(k) \frac{\partial \alpha_i}{\partial w_{6,j,l}} + \sum_{i=1}^n \delta_i(k) w_{5,i,j}(k) \frac{\partial \beta_i}{\partial w_{6,j,l}}. \quad (23)$$

对输入层到隐层的连接权  $w_2$ , 有

$$\Delta w_{2,j,q}(k) = \sum_{i=1}^n \delta_i(k) w_{3,i,j}(k) f_j(\bullet) u_q(k) +$$

$$\sum_{i=1}^n \delta_i(k+1) w_{4,i,j}(k+1) f_j(\bullet) c_j(k) + \sum_{i=1}^n \delta_i(k+2) w_{5,i,j}(k+2) f_j(\bullet) b_j(k). \quad (24)$$

式中  $w_4(k+1)$  已计算出, 其余几个量可按下式近似估算:

$$\delta_i(k+1) = \delta_i(k) + \lambda[\delta_i(k) - \delta_i(k-1)], \quad (25)$$

$$\delta_i(k+2) = \delta_i(k+1) + \lambda[\delta_i(k+1) - \delta_i(k)], \quad (26)$$

$$w_{5,i,j}(k+2) = w_{5,i,j}(k+1) + \eta \Delta w_{5,i,j}(k). \quad (27)$$

### 5 改进型和 PD 型 Elman 网络对动态系统辨识的对比研究

关于神经网络辨识的研究已有许多<sup>[5-7]</sup>, 本文基于图 3 所示的并行动态系统辨识结构, 将改进型和 PD 型 Elman 网络分别用于对动态系统的辨识

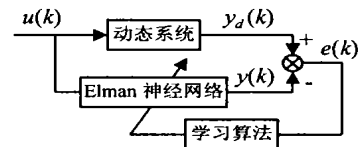


图 3 采用 Elman 网络的动态系统辨识

该结构将输入信号送入神经网络学习, 得到  $y(k)$ ; 将  $y(k)$  和动态系统的实际输出  $y_d(k)$  比较后得到偏差  $e(k)$ . 辨识的目的是令偏差  $e(k) = y_d(k) - y(k)$  尽可能小, 即由学习算法不断调整神经网络内部的权值, 令神经网络的输出不断逼近实际系统的输出. 用于对比研究的非线性动态系统的数学模型为

$$y(k+1) = 0.8y(k) + \frac{1}{y(k)} - 0.5u(k). \quad (28)$$

输入信号  $u(k)$  为两个不同频率的正弦信号相加, 即

$$u(k) = 0.4\sin\left(\frac{3\pi k}{14}\right) + 0.6\sin\left(\frac{2\pi k}{11}\right), \quad (29)$$

其中  $k$  取值为  $1, 2, \dots, 200$ , 即  $u(k)$  和  $y(k)$  组成 200 个数据对.

取隐层、结构层、记忆层均为 15 个节点, 各层均采用 Sigmoid 型函数, 输出层神经元为线性函数, 学习周期为 600 次. 改进型与 PD 型 Elman 网络对式 (28) 的辨识收敛速度比较如图 4 所示, 辨识结果分别见图 5 和图 6.

在网络训练 600 次后, 改进 Elman 的输出均方误差为 0.006133; 而 PD 型 Elman 网络输出的均方误差为 0.001049. 从图 5 和图 6 可以看出, PD 型

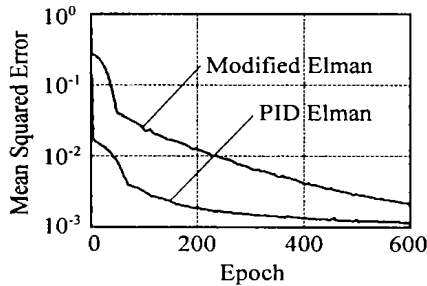


图4 两种 Elman 网络的收敛速度比较

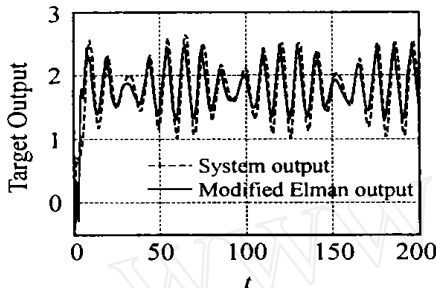


图5 改进 Elman 网络的辨识结果

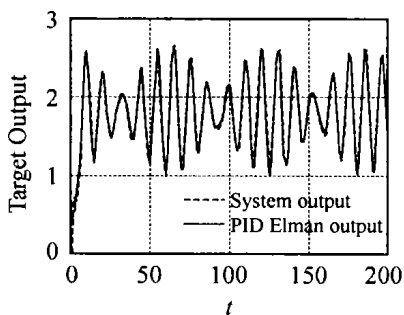


图6 PD 型 Elman 网络的辨识结果

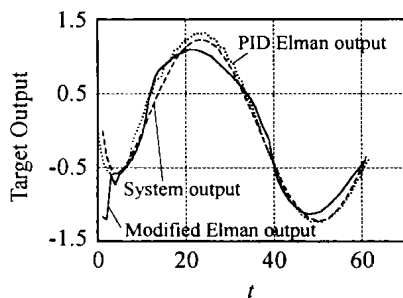


图7 改进型和 PD 型 Elman 网络对电液伺服系统的仿真曲线

Elman 网络的输出非常逼近理想模型的输出,说明它具有更好的动态辨识性能

下面将改进型和 PD 型 Elman 网络用于某炼钢厂热轧机组的液压电液伺服系统。网络结构同前,训练次数增加为 1 000 次。仿真过程中的输入信号和输出信号分别为输入伺服阀的电流信号和液压缸的速度信号,利用学习后的 Elman 网络对液压电液伺服

系统进行辨识,辨识结果如图 7 所示

由图 7 可以看出,改进 Elman 网络辨识初期误差较大,随着时间的增加,逐渐逼近实际系统;而 PD 型 Elman 网络的初始误差明显比改进 Elman 网络的小,且其输出一直能逼近实际系统的输出,具有更好的非线性辨识效果

## 6 结 论

本文在改进 Elman 网络的结构基础上,提出了一种具有 PD 特性的 Elman 网络,并给出了其学习算法。同时搭建了并行结构的动态系统辨识模型,分别针对理想数学模型和实际液压电液伺服系统进行了计算机仿真。仿真结果表明,对于动态非线性系统,PD 型 Elman 网络与普通改进 Elman 网络相比具有更好的辨识性能。该模型对于在线辨识以及模型预测具有较好的参考价值。

## 参考文献 (References)

- [1] Elman J, Finding Structure in Time [J], *Cognitive Science*, 1990, 14(2): 179-211.
- [2] Cheng Y C, Qi W M, Cai W Y. Dynamic Properties of Elman and Modified Elman Neural Network [A]. *Proc of the 1st Int Conf on Machine Learning and Cybernetics* [C]. Beijing, 2002, 2(2): 637-640.
- [3] Ren X M, Chen J, Gong Z H. Approximation Property of the Modified Elman network [J]. *J of Beijing Institute of Technology*, 2002, 11(1): 19-23.
- [4] 程远楚. 水电机组智能控制策略与调速励磁协调控制的研究[D]. 武汉: 华中科技大学, 2002: 125-131. (Cheng Y C. *Study on Intelligent Control Strategies of Hydro-turbine Generating Unit and Coordinated Control for Governing and Excitation Systems* [D]. Wuhan: Huazhong University of Science and Technology, 2002: 121-131.)
- [5] Gao X Z, Gao X M, Ovaska S J. A Modified Elman Neural Network Model with Application to Dynamic Systems Identification [A]. *IEEE Int Conf on Systems, Man, and Cybernetics* [C]. Beijing, 1996: 1376-1381.
- [6] Tan Y H, Dang X J, Liang F, et al. Dynamic Wavelet Neural Network for Nonlinear Dynamic System Identification [A]. *Proc of the 2000 IEEE Int Conf on Control Applications* [C]. Anchorage, Alaska, 2000: 214 - 219.
- [7] Calderon G, Draye J-P, Pavisic D, et al. Nonlinear Dynamic System Identification with Dynamic Recurrent Neural Networks [A]. *Proc of Int Workshop on Neural Networks for Identification, Control, Robotics, and Signal/Image Processing* [C]. Venice, 1996: 49-54.