

文章编号: 1000-8152(2000)01-0092-04

一种基于动态人工神经网络的 Wiener 模型辨识

李世华

吴福保

李奇

(东南大学自动控制系, 南京, 210096) (东南大学自动化研究所, 南京, 210096)

摘要: 提出了一种新的辨识模型对 Wiener 模型进行辨识, 该模型由一线性动态神经元串联一静态 BP 网络模型组成, 利用线性动态神经元对 Wiener 模型的线性动态部分建模, 利用静态 BP 网络逼近模型的静态非线性部分, 并且给出了统一的 BP 辨识算法, 仿真结果表明了该方法的有效性。

关键词: 人工神经网络; 系统辨识; Wiener 模型

文献标识码: A

Identification of Wiener Model Using Dynamic Artificial Neural Networks

LI Shihua

(Department of Automatic Control, Southeast University, Nanjing, 210096, P. R. China)

WU Fubao and LI Qi

(Research Institute of Automation, Southeast University, Nanjing, 210096, P. R. China)

Abstract: The problem of identification of a Wiener model is studied. The proposed identification method uses a dynamic neural network (DANN) which consists of a linear dynamic neuron (LDN) in cascade with a static BP neural network (SBP). A unified back-propagation algorithm is proposed to estimate the weights and the biases of the LDN and the SBP simultaneously. Numerical examples are provided to show the efficiency of the proposed method.

Key words: neural networks; nonlinear systems identification; Wiener model

1 引言 (Introduction)

在实际的系统中, 如 pH 控制、流体控制、生化反应系统等之中, 存在广泛的一类非线性系统可由 Wiener 模型 (下简称 W 模型) 来描述, 该模型由一线性动态子系统与一无记忆非线性增益组成^[1], 其离散化结构如图 1。

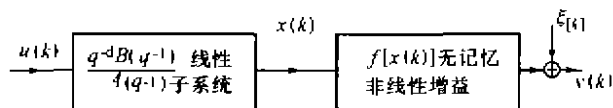


图 1 Wiener 模型
Fig. 1 Wiener model

离散时间 Wiener 模型的差分方程描述为

$$\begin{cases} A(q^{-1})x(k) = q^{-d}B(q^{-1})u(k), \\ y(k) = f[x(k)] + \xi(k). \end{cases} \quad (1)$$

其中

$$A(q^{-1}) = 1 + a_1q^{-1} + \dots + a_nq^{-n}, \quad (2)$$

$$B(q^{-1}) = b_0 + b_1q^{-1} + \dots + b_mq^{-m}, \quad (3)$$

分别为 n, m 阶后移算子多项式, d 为系统的时延, $f(\cdot)$ 为无记忆的非线性增益, $y(k), u(k)$ 和 $\xi(k)$

分别为 k 时刻系统的输出、输入和噪声, $x(k)$ 为 k 时刻线性动态子系统的输出。

有关 W 模型辨识已提出多种, 文[2, 3]采用三电平伪随机序列分别辨识出连续和离散的 W 模型, 文[4]提出采用周期脉冲信号辨识的方法, 这些方法不足之处在于: 采用多项式逼近非线性增益; 且若多项式阶次设置不当会影响辨识精度; 参数估计公式重复计算量大。

近年来, 将人工神经网络应用于系统辨识的研究不少, 也取得了不少成果^[5-7], Narendra 等人在文[8]中提出了一种推广的动态人工神经网络模型及相应的学习算法, 由于要求知道模型中线性子系统传递函数, 因此它的使用受到限制, 文[7]中提出利用 BP 网络单独逼近非线性增益的思想, 采用两步法对 W 模型的两部分分别进行了辨识, 先采用小随机输入信号利用最小二乘法估计出线性动态子系统, 然后利用静态 BP 网络逼近无记忆非线性增益部分, 该方法的局限性在于: 要求无记忆非线性增益部分存在适当的近似线性区域, 并且未考虑噪声情

况;当系统噪声较强时,小随机输入信号激励下对象响应较小,易被噪声信号掩盖而无法进行有效辨识,而增大该小随机输入信号的幅度,易使对线性动态子系统的参数估计偏差增大,导致整体辨识效果变差。

考虑 W 模型的结构特征,本文提出了利用一种动态人工神经网络模型从结构上模拟 W 模型,达到同步辨识 W 模型的效果,而无需分两步进行辨识。同时,因是结构上模拟,辨识中采用的动态人工神经网络结构相对简洁,而且实现算法简单易行,收敛性好。

2 动态人工神经网络模型及算法 (Model and algorithm of dynamic artificial neural networks)

算法要求系统满足下列条件:

- 1) 系统结构参数 n, m, d 已知;
- 2) $f(\cdot)$ 为连续函数;
- 3) $A(q^{-1})$ 为渐近稳定的多项式, $A(q^{-1})$ 和 $B(q^{-1})$ 互质,且 $B(1) \neq 0$ 。

动态人工神经网络(简称 DANN)模型结构见图 2。DANN 结构由一多输入的线性动态神经元(Linear Dynamic Neuron,简称 LDN)串接一单输入单输出的静态 BP(简称 SBP)子网络组成。SBP 为多层前馈网络,此处为了分析讨论方便,只考虑单隐层情况。LDN 输入由系统的输入及 LDN 的输出的时延信号组成,其激励函数为线性。显然, W 模型的线性动态

部分参数可完全由 LDN 的权值来表示。提出这种模型结构的目的,简单来说是利用 LDN 并联模型跟踪 W 模型的线性动态子系统,而 SBP 子网络跟踪 W 模型的非线性增益部分。并且,与其它有关 W 模型的辨识方法相比,利用 LDN 对 W 模型的线性部分来建模使得能够得到统一的 BP 辨识算法,同步估计出 SBP 和 LDN 的权值和阈值。

可以看出, DANN 是从结构上模拟 W 模型的,这相当于在多维搜索空间加上了限制,使权值从一开始就在一个相对较固定的空间划分中搜索最优点,只要权值能最后收敛到最优点,就能够分离出 W 模型的两部分,并且与普通的人工神经网络(简称 ANN)串并联模型^[6](只能跟踪模型输出,无法真正分离模型的两部分)相比,模型的结构相对合理简单,降低了计算量。简单比较一下,设两种模型所需隐层节点数分别为 la, lb ,则对于图 2 模型,总的权值数为 $n + m + 3la + 1$,对于 ANN 串并联模型模型,从输入节点到隐层之间的权值数为 $(n + m + 2) \cdot lb + 1$;而图 2 模型中 la 个隐层节点是用来静态逼近非线性增益,ANN 串并联模型模型中 lb 个隐层节点是用来动态逼近非线性增益的,因此在相同的逼近精度下,两模型所需的隐层数 $la \leq lb$;可以看出,每次训练时,图 2 模型的计算量明显低于 ANN 串并联模型模型的计算量。

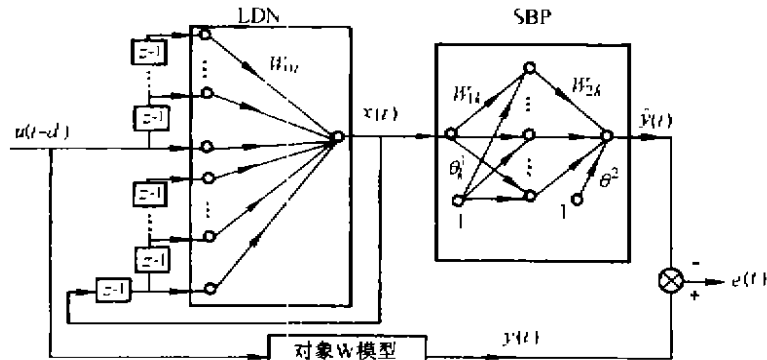


图 2 DANN 并联模型

Fig 2 Parallel connection model of DANN

由 W 模型的特征,有如下性质:当 $\lambda \neq 0$ 时,如下 W 模型与式(1)输入输出等价

$$\begin{cases} A(q^{-1})x(k) = q^{-d}B_1(q^{-1})u(k), \\ y(k) = f_1(x(k)) + \xi(k). \end{cases} \quad (4)$$

其中 $B_1(q^{-1}) = B(q^{-1})/\lambda$, $f_1(x(k)) = f[\lambda x(k)]$, $\xi(k)$ 仍为式(1)所示。

显然,若能利用本文的动态人工神经网络模型同步辨识出 W 模型的两部分 $A(q^{-1})$, $B_1(q^{-1})$ 和 $f_1(x(k))$,则已获得了 W 模型的输入输出等价模型,可以说实现了辨识分离。

模型的算法描述如下:

LDN 子网络输入层为

$$o_i^0(t) = \begin{cases} \hat{x}(t-i), & 1 \leq i \leq n, \\ u(t-d-i+n+1), & n+1 \leq i \leq m+n+1. \end{cases} \quad (5)$$

网络的输出层为

$$\hat{x}(t) = \text{net}^0(t) = \sum_{i=1}^{m+n+1} w_{0i}(t) o_i^0(t), \quad (6)$$

其中, $w_{0i}(t)$ 为权系数。

SBP 子网络输入为:

$$o^1(t) = \text{net}^0(t). \quad (7)$$

隐层为:

$$\text{net}_k^1(t) = w_{1k}(t)o^1(t) + \theta_k^1(t), \quad (8)$$

$$o_k^2(t) = \text{sig}(\text{net}_k^1(t)). \quad (9)$$

其中, $w_{1k}(t)$ 为权值, $\theta_k^1(t)$ 为隐层阈值, $\text{sig}(\cdot)$ 为激励函数, 可取

$$\text{sig}(x) = (1 - e^{-x}) / (1 + e^{-x}). \quad (10)$$

输出层为:

$$\hat{y}(t) = \text{net}^2(t) = \sum_{k=1}^K w_{2k}(t)o_k^2(t) + \theta^2(t). \quad (11)$$

其中, $\hat{y}(t)$ 是 DANN 的输出, $w_{2k}(t)$ 为权值, $\theta^2(t)$ 是阈值, K 为隐层节点数.

利用 BP 算法对 DANN 网络权值与阈值进行学习, 取辨识算法优化指标为

$$J_m = \frac{1}{2} [y(t) - \hat{y}(t)]^2. \quad (12)$$

对 SBP 子网络得到权值及阈值调整公式为:

$$w_{2k}(t+1) = w_{2k}(t) + \eta [y(t) - \hat{y}(t)] o_k^2(t), \quad (13)$$

$$w_{1k}(t+1) = w_{1k}(t) + \eta [y(t) - \hat{y}(t)] \cdot \text{sig}'(\text{net}_k^1(t)) w_{2k}(t) o^1(t), \quad (14)$$

$$\theta^2(t+1) = \theta^2(t) + \eta [y(t) - \hat{y}(t)], \quad (15)$$

$$\theta_k^1(t+1) = \theta_k^1(t) + \eta [y(t) -$$

$$\hat{y}(t)] \text{sig}'(\text{net}_k^1(t)) w_{2k}(t). \quad (16)$$

$$\text{其中 } \text{sig}'(x) = 0.5[1 - \text{sig}^2(x)], \quad (17)$$

η 为学习步长.

对 LDN 子网络:

$$w_{0i}(t+1) = w_{0i}(t) + \eta [y(t) - \hat{y}(t)] o_i^0(t) \cdot \sum_{k=1}^K (w_{2k}(t) \text{sig}'(\text{net}_k^1(t)) w_{1k}(t)). \quad (18)$$

3 仿真结果 (Simulation result)

考虑如下被辨识系统

$$x(t) = 1.4138x(t-1) - 0.6065x(t-2) + 0.1044u(t-1) + 0.0883u(t-2),$$

$$y(t) = \frac{x(t)}{\sqrt{0.10 + 0.90(x(t))^2}} + \xi(t),$$

其中控制输入 u 取 $[-4.0, 4.0]$ 之间的随机数, $\xi(t)$ 为零均值的白噪声, 因此系统是输入输出有界的稳定系统.

由本文方法, 取 LDN 子网络输入节点 4 个, SBP 子网络节点 10 个, 权值阈值学习速率 $\eta = 0.01$. 经过 20000 次的学习, 线性动态子系统的参数估计结果见表 1. 不考虑噪声情况下, 非线性增益的实际输出和辨识输出见图 3.2 万次训练学习之后, 加入测试输入信号, 实际过程的输出和辨识模型的输出见图 4.

表 1 参数估计结果

Table 1 Estimate results of parameter

算法	输入信号	噪声方差	a_1	a_2	b_0/λ	b_1/λ
真值			-1.4138	0.6065	0.1044	0.0883
文[8]	$[-0.05, 0.05]$	0.00	-1.4137	0.6065	0.1044	0.0883
	$[-0.2, 0.2]$	0.01	-1.4265	0.6132	0.1149	0.0834
	$[-0.4, 0.4]$	0.04	-1.3853	0.5709	0.0947	0.0597
	$[-0.6, 0.6]$	0.09	-1.3454	0.5519	0.0870	0.0926
	$[-4.0, 4.0]$	0.00	-1.4139	0.6070	0.1042	0.0885
本文	$[-4.0, 4.0]$	0.01	-1.4165	0.6085	0.1048	0.0880
	$[-4.0, 4.0]$	0.04	-1.4185	0.6116	0.1052	0.0876
	$[-4.0, 4.0]$	0.09	-1.4203	0.6102	0.1052	0.0876

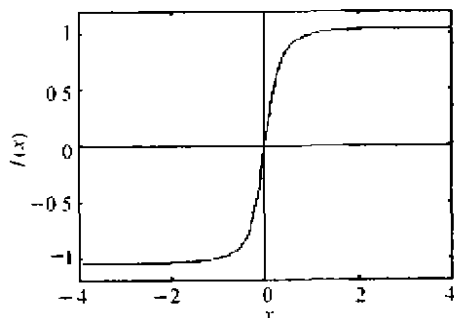


图 3 非线性的实际输出(—)和辨识输出(---)

Fig. 3 Actual and identified output of nonlinearity

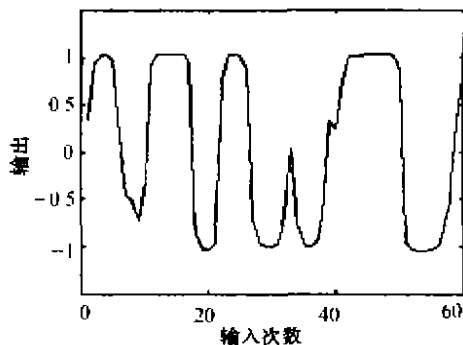


图 4 过程的实际输出(—)和辨识模型的输出(---)

Fig. 4 Output of plant and identified model

采用文[6]中普通 ANN 串并联模型进行辨识, 取 BP 网络结构为 4-20-1, 权值阈值学习速率 $\eta = 0.01$, 经过 100000 次的学习, 辨识结果仍有差异, 两种方法的辨识均方差曲线比较见图 5。

可见, 对于 W 模型的辨识, 文[6]的模型虽然跟踪速度较快, 权值能短时间内收敛到真值的附近, 但经过长时间训练仍存在较大差异; 而本文方法虽然开始时跟踪较慢, 但经过一段时间训练后, 权值能最终收敛到真值, 这是因为本文模型是从结构上模拟 W 模型的。

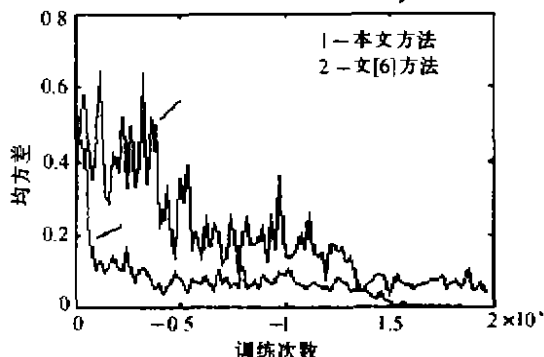


图 5 辨识均方差曲线比较
Fig. 5 Error curve

4 结论 (Conclusion)

仿真结果说明了该方法的有效性, 在噪声情况下, 仍能获得较高精度的估计结果, 并且该方法模型结构简单, 计算量少, 收敛性较好。

(上接第 91 页)

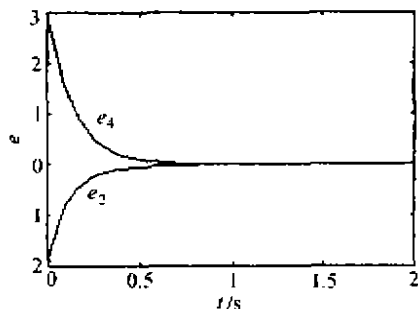


图 1 状态估计误差 e_2, e_4
Fig. 1 Estimated state errors e_2, e_4

观测误差收敛速度, 这正是我们给出的理论分析结果。

参考文献 (References)

- [1] Walcott B L, Corless M J and Zak S H. Comparative study of nonlinear state-observation techniques [J]. Int. J. Control, 1987, 45 (6): 2109 - 2132
- [2] Raghavan S and Hedrick J K. Observer design for a class of nonlinear systems [J]. Int. J. Control, 1994, 59 (2): 515 - 528
- [3] 马克茂, 王子才, 史小平. 一类非线性系统的观测器设计 [J]. 控

参考文献 (References)

- [1] Hsia T C 著, 关礼民译. 系统辨识与应用 [M]. 长沙: 中南工业大学出版社, 1986
- [2] Krempel R. Application of three-level-pseudorandom-signals for parameter estimation of nonlinear system [A]. 3rd IFAC Symp. on Identification and System Parameter Estimation [C], 1973, 2: 835 - 838
- [3] 胡德文, 王正志. 非线性系统 Wiener 模型辨识 [J]. 自动化学报, 1991, 17 (2): 151 - 159
- [4] 黄正良, 吴坚, 万百五. 辨识 Wiener 模型的一种新方法 [J]. 控制理论与应用, 1996, 13 (3): 326 - 332
- [5] Chen S, et al. Nonlinear system identification using neural networks [J]. Int. J. Control, 1990, 51 (6): 1191 - 1214
- [6] 田明, 戴汝为. 基于动态 BP 神经网络的系统辨识方法 [J]. 自动化学报, 1993, 19 (4): 450 - 453
- [7] Al-Duwaihi H, et al. Use of multilayer feedforward neural networks in identification and control of Wiener model [J]. IEE Proc. Control Theory and Application, 1996, 143 (3): 25 - 258
- [8] Narendra K S and Parthasarathy K. Identification and control of dynamic systems using neural networks [J]. IEEE Trans Neural Networks, 1990, 1 (1): 4 - 27

本文作者简介

李世华 1975 年生, 分别于 1995 年, 1998 年在东南大学获得学士和硕士学位, 现于东南大学自动控制系攻读博士学位, 主要研究方向为: 系统辨识, 智能控制, 机器人控制。

吴福保 1972 年生, 1997 年在东南大学获得硕士学位, 现于东南大学自动化所攻读博士学位, 主要研究方向为: 粗糙集合理论与应用, 模糊控制, 神经网络控制。

李 奇 1963 年生, 1992 年在东南大学获博士学位, 现为东南大学教授, 主要研究方向: 系统辨识, 复杂系统智能控制与计算机控制。

制理论与应用, 1998, 15 (3): 443 - 446

- [4] Ciccarella G, Dalla Mora M and Germani A. A luenberger-like observer for nonlinear systems [J]. Int. J. Control, 1993, 57 (3): 537 - 556
- [5] Cheng D Z and Yun X P. Linearization of time-varying affine nonlinear systems [J]. System Sciences and Mathematical Sciences, 1988, 1 (2): 119 - 130
- [6] Slotine J J E and Li W. 应用非线性控制 [M]. 蔡自兴等译. 北京: 国防工业出版社, 1992

本文作者简介

向峰嵘 1969 年生, 讲师, 1998 年在南京理工大学获博士学位, 主要研究兴趣是非线性控制, 不确定性系统的鲁棒控制, 神经网络在控制中的应用等。

陈庆伟 1963 年生, 高级工程师, 1988 年在南京理工大学自动控制系获硕士学位, 目前主要研究方向为智能控制理论及应用, 计算机控制系统等。

吴晓蕾 1958 年生, 副教授, 1986 年在南京理工大学自动控制系获硕士学位, 目前主要研究方向为智能控制理论及神经网络在控制中的应用等。

胡维礼 1941 年生, 南京理工大学教授, 博士生导师, 1965 年毕业于清华大学自控系, 目前研究领域为鲁棒自适应控制, 非线性控制及智能控制理论在伺服驱动系统中的应用等。