第 26 卷第 11 期
2009 年 11 月

控 制 理 论 与 应 用
Control Theory & Applications

Vol. 26 No. 11
Nov. 2009

文章编号: 1000−8152(2009)11−1192−05

非线性动态系统的Wiener神经网络辨识法

吴德会

(九江学院 数字控制技术与应用江西省重点实验室, 江西 九江 332005; 清华大学 电力系统国家重点实验室, 北京 100084)

摘要: 提出了一种新的Wiener神经网络结构并将其应用于非线性动态系统辨识问题. 首先, 用Wiener模型对非线性动态系统进行描述, 将其分解成线性动态子环节串接非线性静态增益的形式. 其次, 设计一种新型的神经网络结构, 使网络权值对应于相应的Wiener模型参数；并推导了基于反向传播的网络权值调整方法. 最后, 通过网络迭代训练, 可同时得到线性动态子环节和非线性静态增益的模型参数. 通过一个Wiener模型的数值仿真来验证方法的有效性, 仿真结果表明所提辨识方法切实可行.

关键词: 非线性动态系统; 辨识; 神经网络; Wiener模型

中图分类号: TP271　　　**文献标识码**: A

Identification method for nonlinear dynamic system using Wiener neural network

WU De-hui

(Key Laboratory of Numerical Control of Jiangxi Province, Jiujiang University, Jiujiang Jiangxi 332005, China;
State Key Laboratory of Power Systems, Department of Electrical Engineering, Tsinghua University, Beijing 100080, China)

Abstract: A novel Wiener neural network structure is presented and applied to nonlinear dynamic system identification. Firstly, the nonlinear dynamic system is described by a Wiener model which consists of a linear dynamic part in cascade with a nonlinear static gain. Secondly, a novel neural network structure is designed, the weights in which are corresponding with the parameters of the Wiener model. Thirdly, backward-propagation methods for the adjustment of weights in the network are discussed. Finally, parameters of the linear dynamic part and the nonlinear static gain in the Wiener model are determined simultaneously by iterative training. A numerical simulation of Wiener model is provided to validate the effectiveness. Simulation results show that the suggested identification schemes are practically feasible.

Key words: nonlinear dynamic system; identification; neural network; Wiener model

1 引言(Introduction)

非线性动态系统辨识是系统辨识领域研究的难点和焦点问题[1,2]. 目前, 以人工神经网络(artificial neural network, ANN)为主要代表的人工智能技术被广泛应用于非线性动态系统辨识[3]. 如Narendr等提出的基于反向传播神经网络的并行和串-并行非线性辨识方法[4]; 吕立华用改进的多分辨小波网络辨识非线性动态系统[5]; 胡玉玲等在模糊神经网络基础上, 形成可将暂态信息记忆于网络的动态回归层, 实现了对非线性动态系统的辨识[6]. 李鸿儒等又将递归神经网络的递推预报误差(recursive prediction error, PRE)算法应用于非线性动态系统建模[7]. 但是从本质上来说, 该类方法是一种函数逼近, 实际上类似一个黑箱, 缺少透明度.

Hammerstein和Wiener等级联模型可以描述非常广泛的一类非线性系统, 通过串联线性动态环节和非线性静态增益, 从结构上就直观地描述了该类系统特性[8]. 我国学者在此辨识领域进行了卓有成效地研究, 胡德文等以不同幅值三电平伪随机m序列为输入, 成功辨识了Wiener模型线性脉冲响应函数和非线性增益系数[9]. 文献[10]中郎自强提出一种开环辨识Hammerstein模型的新方法, 可获得模型线性动态部分参数的渐近无偏估计. 黄正良提出辨识联级模型(Hammerstein和Wiener)的两步法[11], 分步骤辨识联级模型的静态非线性增益和线性动态环节, 并已取得较好效果, 但该辨识过程相对繁琐. 张艳等在文献[12]中提出基于粒子群优化的Wiener模型辨识新方法, 但由于实际中Wiener模型的非线性增益

收稿日期: 2008−05−29; 收修改稿日期: 2009−01−05.
基金项目: 国家自然科学基金资助项目(50705039); 中国博士后基金资助项目(20070420358).

是未知的, 使该方法的应用受到限制.

本文作者在文献[13]提出了一种利用支持向量回归机(support vector regression, SVR)的非线性动态系统辨识方法, 但是从理论上不难证明, 该方法存在无穷多解, 为此, 文献[13]中增加了约束条件, 使得满足约束条件的解唯一(特解). 虽然该特解能够满足模型方程, 但由于额外增加了约束条件, 因此无法保证所求的该特解是所有泛解中较优的.

本文在对Wiener非线性动态模型结构认识基础上, 定义优化目标函数, 提出一种Wiener神经网络结构. 推导目标函数与模型各参数之间导数关系, 由此给出网络训练算法, 实现对Wiener模型中的线性动态环节和非线性增益的同时辨识. 该方法能给出Wiener模型的数学解析表达, 且辨识结果唯一.

2　非线性问题描述(Nonlinear problem formulation)

Wiener模型可看成是一个线性动态环节与一个非线性静态增益的串联组合. 对于单输入单输出(single-input and single-output, SISO)的离散时间Wiener模型如图1所示[8], 其差分方程描述如下:

$$\begin{cases} A(q^{-1})x(t) = q^{-d}B(q^{-1}), \\ y(t) = f[x(t)] + e(t). \end{cases} \tag{1}$$

其中: $A(q^{-1}) = 1 + a_1q^{-1} + \cdots + a_nq^{-n}$, $B(q^{-1}) = b_0 + b_1q^{-1} + \cdots + b_mq^{-m}$分别为$n$和$m$阶后移算子多项式, 有$q^{-i}u(t) = u(t-i)$成立, d为系统时延, $u(t)$和$y(t)$分别是系统的输入和输出, $e(t)$为系统干扰, $x(t)$是中间信号, 既是线性动态环节的输出, 又是非线性静态增益的输入, 在实际过程中是不可测量的, $f(t)$为非线性静态增益函数.

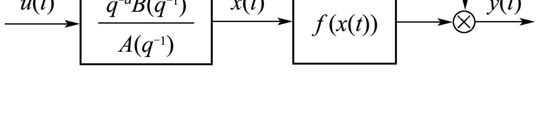

图 1　Wiener非线性动态模型

Fig. 1　Wiener nonlinear dynamic model

非线性静态增益函数$f(\cdot)$通常可用p次多项式近似表达, 则Wiener模型输出为

$$y(t) = c_1x(t) + c_2x^2(t) + \cdots + c_px^p(t) + e(t). \tag{2}$$

Wiener模型是参数模型, 因此可定义新的参数向量$\theta = [a_1, \cdots, a_n, b_0, b_1, \cdots, b_m, c_1, \cdots, c_p]^{\mathrm{T}}$. 需要强调的是, Wiener模型的参数解$\theta$并不是唯一的, 对于任意非零$k$, 由$k(\dfrac{q^{-d}B(q^{-1})}{A(q^{-1})})$和$f(\dfrac{x(t)}{k})$描述的Wiener模型具有相同的输入、输出特性, 因此, 理想的Wiener模型参数θ存在无穷多解. 考虑到观测时的噪声干扰, 实际系统的Wiener模型辨识结果会存在无穷多近似解. 文献[13]中, 假设动态环节终态增益为1, 即通过增加约束条件$\sum\limits_{i=0}^{m} b_i - \sum\limits_{i=1}^{n} a_i = 1$, 使模型辨识结果唯一, 但不能保证该特解一定优于其他近似解.

记$y'(t)$为第t时刻激励$u(t)$作用下的Wiener模型输出, 偏差$\varepsilon(t) = y'(t) - y(t)$ 表示模型在第t个样本上的误差, 则存在误差序列$\{\varepsilon(t)\}_{t=1}^{N}$且该序列的分布情况直接体现了Wiener模型的建模精度. 因此, 模型"准确性"和"稳定性"可分别通过误差序列的均值和均方差指标来描述, 即有

$$M = \frac{1}{N}\sum_{i=1}^{N}|\varepsilon(t)|, \tag{3}$$

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}[\varepsilon(t) - M]^2}. \tag{4}$$

M和σ能综合体现建模误差分布情况, 其值越小, 模型对实际问题逼近越好. 但优化指标含有M和σ两个目标, 是个多目标优化问题, 直接计算量较大. 本文引入均方误差(mean square error, MSE)解决:

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}\varepsilon^2(t) = (M^2 + \sigma^2). \tag{5}$$

由式(5)可知, MSE与$M^2 + \sigma^2$能反映相同的分布信息, 若以MSE作为评价指标可将多目标优化转换为单目标问题. 为计算方便, 可在MSE前乘上系数$\dfrac{1}{2}$, 给出了如下模型辨识目标函数的定义:

定义 1　对于非线性系统输入、输出观测序列$\{u(t), y(t)\}_{t=1}^{N}$, N 为观测窗口长度, 则Wiener模型辨识的目标评价函数为$J = \dfrac{1}{2N}\sum\limits_{i=1}^{N}[y(t) - y'(t)]^2$, 其中$y'(t)$为Wiener模型输出.

由定义1, 最优Wiener模型参数θ应满足下式:

$$\min_{\theta} J = \min_{\theta}\frac{1}{2N}\sum_{i=1}^{N}[y(t) - y'(t)]^2. \tag{6}$$

3　Wiener神经网络及其训练算法(Wiener neural network and training algorithm)

3.1　Wiener神经网络结构的提出(Proposal of Wiener neural network)

本文利用函数连接型神经网络(functional link ANN, FLANN)思想[14], 对系统输入$u(t)$和中间信号$x(t)$分别进行时延和幂级数展开, 以构造了一种特殊的人工神经网络结构. 在该网络结构中, 每一个神经网络权值与实际的Wiener模型参数相对

应, 如此, 非线性Wiener模型辨识问题就可以通过对神经网络的训练来实现. 在图1所示的SISO系统中, Wiener模型中线性动态环节又可用下式来描述:

$$x(t) = (-a_1q^{-1} - a_2q^{-2} - \cdots - a_nq^{-n})x(t) + (b_0q^{-d}+b_1q^{-d-1}+\cdots+b_mq^{-d-m})u(t). \quad (7)$$

再结合由p次多项式近似表达的非线性增益式(2), 本文设计相应多层神经网络结构如图2所示.

由图2可以看出, 在神经网络的输入层, 利用延迟算子q^{-1}将系统输入信号扩展成具有时延特性的系列节点$u(t-d), u(t-d-1), \cdots, u(t-d-m)$, 再结合隐含层的反馈信息$x'(t)$来表达Wiener模型中的线性动态环节. 在输出层, 通过幂级数函数扩展表达对隐含层信号$x'(t)$的非线性映射关系, 与Wiener模型中的非线性静态增益环节相对应. 则Wiener神经网络输出层的结果可表达成

$$y'(t) = f_{\mathrm{out}}(c_1, c_2, \cdots, c_p, x'(t)) = \sum_{i=1}^{p} c_i x'^{i}(t). \quad (8)$$

隐含层节点$x'(t)$为

$$x'(t) = -\sum_{i=1}^{n} a_i x'(t-i) + \sum_{i=0}^{m} b_i u(t-i), \quad (9)$$

其中: 网络权值$a_1, a_2, \cdots, a_n, b_0, b_1, \cdots, b_m$即为Wiener模型线性动态环节参数的估计值; $c_1, c_2, \cdots, c_p$为非线性增益参数的估计值. 由此看来, 通过这种特殊神经网络结构, 可将网络权值与待辨识Wiener模型参数等价起来, θ既是网络的权向量, 也是模型的参数向量. 从而为通过神经网络训练, 达到Wiener模型参数辨识提供了一种新的尝试.

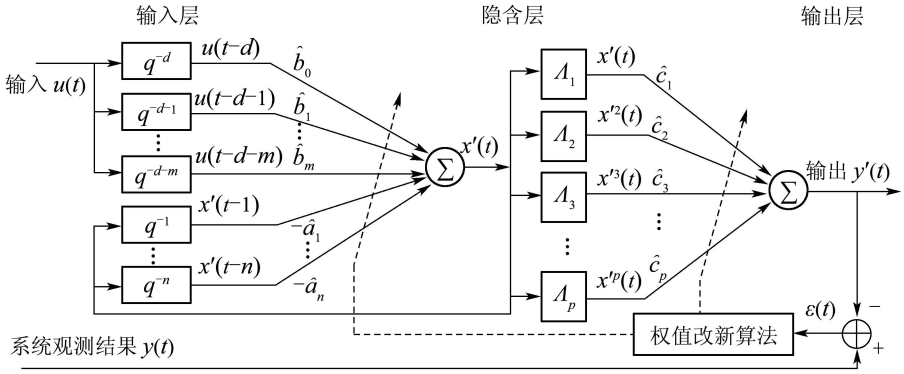

图 2 Wiener神经网络结构

Fig. 2 Wiener neural network structure

3.2 优化目标的梯度计算(Gradient calculation of the optimization objective)

以定义1确定的函数J为模型辨识优化目标, 采用负梯度下降学习原理对Wiener神经网络参数θ进行更新. 根据目标函数J的定义, 可推导其梯度为

$$G = \frac{\partial J}{\partial \theta} = \sum_{i=1}^{N}(\varepsilon(t)\frac{\partial \varepsilon(t)}{\partial \theta}). \quad (10)$$

将偏差含义$\varepsilon(t) = y'(t) - y(t)$代入上式, 得

$$G = \sum_{i=1}^{N}(\varepsilon(t)\frac{\partial y'(t)}{\partial \theta}). \quad (11)$$

分别对参数向量θ 3组参数的偏导数进行推导. 首先, 由Wiener神经网络输出层表达式(8)计算$y'(t)$对隐含层权值参数的偏导为

$$\frac{\partial y'(t)}{\partial c_i} = x'^{i}(t),\ i = 1, \cdots, p \quad (12)$$

及对隐含层节点$x'(t)$的偏导为

$$\frac{\partial y'(t)}{\partial x'(t)} = i\sum_{i=1}^{p} c_i x'^{i-1}(t),\ i = 1, \cdots, p. \quad (13)$$

对于隐含层节点$x'(t)$, 有如下的递推关系:

$$x'(t) = -\sum_{i=1}^{n} a_i x'(t-i) + \sum_{i=0}^{m} b_i u(t-i). \quad (14)$$

由此可见, 其对a_i, b_i偏导为

$$\frac{\partial x'(t)}{\partial a_i} = -x'(t-i) - \sum_{j=1}^{n} a_j \frac{\partial x'(t-j)}{\partial a_i}, \quad (15)$$

$$\frac{\partial x'(t)}{\partial b_i} = u(t-i) - \sum_{j=1}^{n} a_j \frac{\partial x'(t-j)}{\partial b_i}. \quad (16)$$

可由复合函数求导原理, 计算网络输出$y'(t)$对输入层权值参数的偏导:

$$\frac{\partial y'(t)}{\partial a_i} = \frac{\partial y'(t)}{\partial x'(t)}\frac{\partial x'(t)}{\partial a_i},\ i = 1, \cdots, n. \quad (17)$$

$$\frac{\partial y'(t)}{\partial b_i} = \frac{\partial y'(t)}{\partial x'(t)}\frac{\partial x'(t)}{\partial b_i},\ i=1,\cdots,m. \quad (18)$$

3.3 Wiener神经网络训练算法(Training algorithm of Wiener neural network)

对于某一确定的SISO非线性系统输入输出观测序列$\{u(t),y(t)\}_{t=1}^{N}$, 可通过负梯度下降法进行训练. 根据式(11), 其权向量迭代过程可表达为

$$\theta(k+1)=\theta(k)+\Delta\theta(k). \quad (19)$$

其中: $\theta(k+1)$为第k轮训练中权向量θ取值, $\Delta\theta(k)$为更新量. 若将每一轮训练过程中的更新量$\Delta\theta(k)$分解到每步迭代中去, 则可得迭代过程中每步权值参数的更新量$\Delta\theta$为

$$\Delta\theta=-\eta\varepsilon(t)\frac{\partial y'(t)}{\partial\theta}. \quad (20)$$

由此, 可得输入层权值参数调整表达式:

$$\Delta a_i=-\eta\varepsilon(t)i\sum_{i=1}^{p}c_i x'^{i-1}(t)[-x'(t-i)-\sum_{j=1}^{n}a_j\frac{\partial x'(t-j)}{\partial a_i}],\ i=1,2,\cdots,n, \quad (21)$$

$$\Delta b_i=-\eta\varepsilon(t)i\sum_{i=1}^{p}c_i x'^{i-1}(t)[-u(t-i)-\sum_{j=1}^{n}a_j\frac{\partial x'(t-j)}{\partial b_i}],\ i=1,2,\cdots,m. \quad (22)$$

隐含层权值参数调整表达式:

$$\Delta c_i=-\eta\varepsilon(t)x'^{i}(t),\ i=1,2,\cdots,p, \quad (23)$$

其中$\Delta a_i,\Delta b_i,\Delta c_i$分别为Wiener神经网络权值参数在当前时刻的调整量.

在每一轮训练开始时, 式(21)和式(22)中相关参数的初值可置为零. 迭代过程中, 可通过设置固定次数或设置目标函数指标的减小量阈值, 作为网络训练停止的条件, 最终的网络权值向量θ即为待辨识Wiener模型参数.

4 数值仿真与分析(Numerical simulation and analyses)

为验证本文方法, 考虑文献[2]中测试用的非线性动态系统的线性动态环节和非线性静态增益

$$\begin{cases} x(t)=0.3x(t-1)+0.6x(t-2)+u(t)+\varepsilon(t),\\ y(t)=x^3(t)+0.3x^2(t)-0.4x(t)+e(t). \end{cases}$$

其中: 噪声$\varepsilon(t),e(t)$选用方差为系统输出强度0.01的零均值Gauss白噪声, 激励信号$u(t)$为

$$u(t)=\frac{1}{20}[\sin(\frac{2\pi}{250}t)+\sin(\frac{2\pi}{25}t)],$$

其取样长度为500. 激励信号$u(t)$和非线性动态系统的观测值(含噪声)$y(t)$如图3所示.

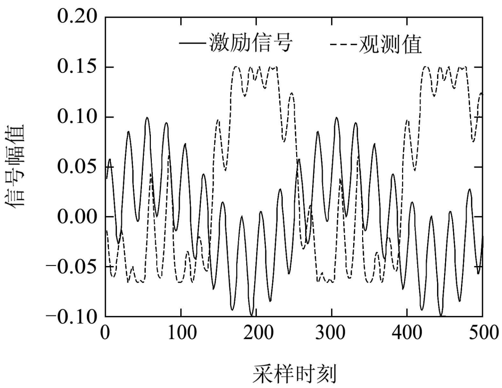

图 3　非线性动态系统输入输出序列

Fig. 3　Input and output sequences of nonlinear system

将系统激励$u(t)$和观测值$y(t)$输入如图2所示的Wiener神经网络进行训练. 由于系统动态环节的阶次m,n, 及幂级数次数p事先未知, 在实际训练时, 可凭借经验来设置. 本实验中, 网络的后移算子n和m均取2, 隐含层节点数p取3, 系统时延$d=0$, 网络的学习因子$\eta=1$. 本例中, Wiener神经网络训练的MSE随迭代次数的变化如图4所示.

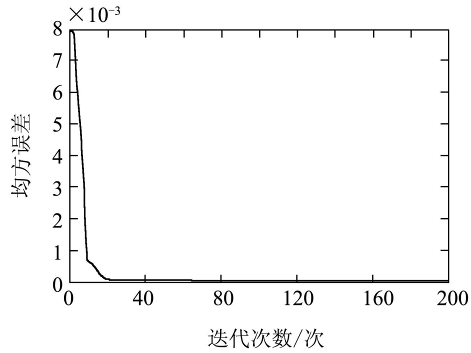

图 4　非Wiener神经网络收敛过程

Fig. 4　Convergence process of Wiener neural network

如图4所示, 在训练初期, Wiener神经网络收敛非常迅速, 迭代到20步之后, 网络就已基本收敛, 输出MSE小于0.0001. 本文给出Wiener神经网络迭代100次之后的模型辨识结果为

$$\begin{aligned} x(t)=\ &0.3444x(t-1)+0.5506x(t-2)-\\ &0.8127u(t)-0.3752u(t-1)+\\ &0.0006u(t-2),\\ y(t)=\ &-0.6748x^3(t)+0.2203x^2(t)+\\ &0.3520x(t). \end{aligned}$$

辨识模型输出$y'(t)$与原系统观测值$y(t)$之间的MSE为9.7×10^{-6}. 虽然辨识结果与原模型在表达式上存在明显的差距, 但模型输出与原系统观测值之间吻合良好. 若继续进行训练, 则可在诸多近似解中逐渐逼近满足式(6)条件的最优解. 但该过程较缓慢, 就本例问题, 迭代10000次之后, 网络MSE才不再减小, 最终辨识结果为

$$\begin{aligned} x(t) = {} & 0.2943x(t-1)+0.6925x(t-2)+ \\ & 1.0044u(t)-0.0083u(t-1)+ \\ & 0.0423u(t-2), \\ y(t) = {} & 1.2720x^3(t)+0.3516x^2(t)-0.4336x(t). \end{aligned}$$

此时, 网络的$\mathrm{MSE_{optimal}}=2.1562\times10^{-6}$. 就解析表达式来看, 辨识模型与原模型并不完全相同. 计算原模型输出与系统观测值$y(t)$之间的$\mathrm{MSE_{original}}=2.8354\times10^{-6}$. 可见, 对于确定参数非线性系统在噪声干扰下, 等效模型参数将发生偏差. 在本例中, 原模型理想输出的误差反而更大, 即$\mathrm{MSE_{original}}>\mathrm{MSE_{optimal}}$，表明本文所提方法辨识的Wiener模型结果优于原系统模型.

当输入信号幅度小于1时, Wiener神经网络具有较好的收敛性; 但当激励远大于1时, 多项式中高阶项会增加拟合误差, 导致网络训练过程的不稳定. 因此, 对于这种情况, 需要采用量纲调整或归一化进行预处理, 以提高网络的可靠性.

5 结论(Conclusions)

1) 非线性动态系统Wiener模型辨识, 存在无穷多近似解, 虽在表达形式上存在明显不同, 但却都具有相似的输入、输出特性.

2) Wiener神经网络以MSE为优化目标, 可实现对线性动态环节和非线性增益的同时辨识, 并得到数学解析表达.

3) Wiener神经网络在不增加约束条件下, 能产生唯一的辨识结果, 在一定程度上克服了Wiener模型辨识中的无穷多解问题.

4) Wiener神经网络虽可方便地得到一个近似数据描述模型. 但仍不可避免会引入传统神经网络方法中局部极小问题, 可利用变学习率法、冲量法、Vogl快速算法等进行改进以提高网络训练性能, 对此还有待进一步研究.

参考文献(References):

[1] KOUKOULAS P, KALOUPTSIDIS N. Nonlinear system identification using Gaussian inputs[J]. *IEEE Transactions on Signal Processing*, 1995, 43(8): 1831 – 1841.

[2] 冯培悌. 系统辨识[M]. 杭州: 浙江大学出版社, 1999.

[3] LEVIN A U, NARENDRA K S. Control of nonlinear dynamical systems using neural networks - Part II: observability, identification, and control[J]. *IEEE Transactions on Neural Networks*, 1996, 7(1): 30 – 42.

[4] NARENDRA K S, PARTHASARATHY K. Identification and control of dynamical systems using neural networks[J]. *IEEE Transactions on Neural Networks*, 1990, 1(1): 4 – 27.

[5] 吕立华, 宋执环, 李平. 用改进的多分辨小波网络辨识非线性动态系统[J]. 浙江大学学报, 2002, 36(1): 36 – 39.
(LÜ Lihua, SONG Zhihuan, LI Ping. Identifying nonlinear dynamic system using the improved multiresolution wave-net[J]. *Journal of Zhejiang University*, 2002, 36(1): 36 – 39.)

[6] 胡玉玲, 曹建国. 基于模糊神经网络的动态非线性系统辨识研究[J]. 系统仿真学报, 2007, 19(3): 560 – 562.
(HU Yuling, CAO Jianguo. Research on identification of dynamic nonlinear system based on fuzzy neural network[J]. *Journal of System Simulation*, 2007, 19(3): 560 – 562.)

[7] 李鸿儒, 顾树生, 邓长辉. 递归神经网络的RPE算法及其在非线性动态系统建模中的应用[J]. 东北大学学报, 2000, 21(6): 590 – 593.
(LI Hongru, GU Shusheng, DENG Changhui. Recursive prediction error algorithm of recurrent neural networks and its application on nonlinear dynamic system modelling[J]. *Journal of Northeastern University*, 2000, 21(6): 590 – 593.)

[8] NARENDRA K S, GALLMAN P G. An interative method for the identification of nonlinear systems using a Hammerstein model[J]. *IEEE Transactions on Automatic Control*, 1966, 11(6): 546 – 550.

[9] 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.
(HU Dewen, WANG Zhengzhi. An identification method for the Wiener model of nonlinear systems[J]. *Acta Automatica Sinica*, 1991, 17(2): 151 – 159.)

[10] 郎自强. 一种辨识Hammerstein模型的新方法[J]. 自动化学报, 1993, 19(1): 37 – 45.
(LANG Ziqiang. A new method for the identification of Hammerstein model[J]. *Acta Automatica Sinica*, 1993, 19(1): 37 – 45.)

[11] 黄正良, 万百五, 韩崇昭. 辨识Hammerstein模型的两步法[J]. 控制理论与应用, 1995, 12(1): 34 – 39.
(HUANG Zhengliang, WAN Baiwu, HAN Chongzhao. A two-stage identification technique for Hammerstein model[J]. *Control Theory & Applications*, 1995, 12(1): 34 – 39.)

[12] 张艳, 李少远, 王笑波, 等. 基于粒子群优化的Wiener模型辨识与实例研究[J]. 控制理论与应用, 2006, 23(6): 991 – 995.
(ZHANG Yan, LI Shaoyuan, WANG Xiaobo, et al. Particle swarm optimal identification of Wiener model and a case study[J]. *Control Theory & Applications*, 2006, 23(6): 991 – 995.)

[13] 吴德会. 基于支持向量机的非线性动态系统辨识方法[J]. 系统仿真学报, 2007, 19(14): 3169 – 3171.
(WU Dehui. Identification for nonlinear dynamic system based on SVM[J]. *Journal of System Simulation*, 2007, 19(14): 3169 – 3171.)

[14] PATRA J C, PANDA G, BALIARSINGH R. Artificial neural network-based nonlinearity estimation of pressure sensors[J]. *IEEE Transactions on Instrumentation and Measurement*, 1994, 43(6): 874 – 881.

作者简介:

吴德会 (1975—), 男, 清华大学电机工程与应用电子技术系博士后, 副教授, 目前研究方向为智能测试与智能控制, E-mail: wudehui@tsinghua.edu.cn.