

## Hammerstein-Wiener 时变系统的带遗忘因子学习辨识算法<sup>①</sup>

仲国民<sup>②</sup> 俞其乐 汪黎明

(浙江工业大学信息工程学院 杭州 310023)

**摘要** 针对一类有限区间上重复运行的 Hammerstein-Wiener 非线性时变系统,将 Hammerstein-Wiener 系统输出非线性部分进行多项式展开以构造回归模型,采用带遗忘因子迭代学习梯度算法和带遗忘因子迭代学习最小二乘法,估计系统的时变参数。当系统参数沿时间轴快变、沿迭代轴缓变时,修正遗忘因子提高算法的辨识精度。文中分别给出了 2 种算法的推导过程并进行仿真对比验证,结果表明,带遗忘因子迭代学习最小二乘法收敛速度更快、精度更高,达到相同性能指标时所需迭代次数更少,验证了所提学习算法的有效性。

**关键词** 学习辨识;最小二乘;随机梯度;Hammerstein-Wiener 模型

### 0 引言

时变参数估计问题在系统辨识领域一直受到人们关注。实际场合中广泛存在时变性和非线性,因此时变非线性系统的辨识问题逐渐成为系统辨识领域较为热门的研究课题之一<sup>[1-3]</sup>。

线性系统的模型结构可以用统一的模型描述,如差分方程、状态空间模型等<sup>[4]</sup>,非线性系统结构复杂,难以直接进行参数估计。基于块结构的非线性模型在诸如化学过程和生物生理系统<sup>[3]</sup>、热反应和化学反应<sup>[5]</sup>、控制系统<sup>[6]</sup>和生物信息学<sup>[7]</sup>中被广泛采用。本文研究的 Hammerstein-Wiener 系统<sup>[1]</sup>作为块结构非线性系统中最为复杂的一种,该模型由输入静态非线性环节、动态线性子系统以及输出静态非线性环节串联构成,通常用来描述在输入输出部分存在非线性元素的系统。现有的块结构 Hammerstein-Wiener 非线性系统参数估计方法主要有:最优两阶段辨识法<sup>[1]</sup>、盲辨识法<sup>[2]</sup>、迭代辨识法<sup>[8]</sup>、偏差补偿最小二乘法<sup>[9]</sup>、递推辨识<sup>[10]</sup>和梯度或最小

二乘迭代辨识法<sup>[11]</sup>等。文献<sup>[12]</sup>借助线性变权粒子群算法对 Hammerstein-Wiener 时滞系统进行辨识,利用粒子群优化和迭代辨识技术实现同步控制所有参数和未知时延的估计。文献<sup>[13]</sup>突破了 Hammerstein-Wiener 非线性系统元素参数化的限制,线性块参数估计可以在一个阶段进行。文献<sup>[14]</sup>针对 Wiener-Hammerstein 系统辨识问题,提出了 2 种估计分别恢复静态非线性特性和线性动态块。文献<sup>[15]</sup>使用非参数高斯过程对结构化非线性系统进行建模,并验证了所提出的非参数建模方法可以有效地建模和辨识。

针对特定模型的非线性系统,需要将现有的先验信息,特别是结构先验知识纳入到算法中,通常会采用特定的辨识算法。现有的递推最小二乘法被证明不具备时变参数的跟踪能力,本文提出的迭代学习辨识方法,其动机来自迭代学习控制的基本思想<sup>[16]</sup>。对于递推算法而言,假设时变参数的变化规律未知,其估计就难以获得一致收敛性。针对重复运行的系统,无论其时变参数随时间如何变化,沿任意时刻重复轴对应的参数缓变,考虑将时间轴上的

<sup>①</sup> 国家自然科学基金(62073291)资助项目。

<sup>②</sup> 男,1983 年生,硕士,讲师;研究方向:系统辨识与学习控制;联系人,E-mail: zgm@zjut.edu.cn。  
(收稿日期:2021-12-23)

递推思想运用于重复轴的学习,构建沿重复轴的“递推”算法<sup>[17]</sup>。针对文献[18]中所给有限区间上重复运行的 Hammerstein-Wiener 非线性时变系统,非线性输出部分是可逆的,将该部分反函数通过多项式展开以构造回归模型。给出能够衡量重复估计性能的指标函数,推导该系统带遗忘因子迭代学习梯度算法(iterative learning gradient algorithm with forgetting factor for Hammerstein-Wiener systems, HW-FILG)和带遗忘因子迭代学习最小二乘算法(iterative learning least squares algorithm with forgetting factor for Hammerstein-Wiener systems, HW-FILLS)。仿真结果显示在重复持续激励条件下,2种算法能够获得时变参数的完全估计,验证了该算法的有效性。

## 1 问题描述

考虑如下时变 Hammerstein-Wiener 非线性系统,如图 1 所示,其线性动态子系统描述如下。

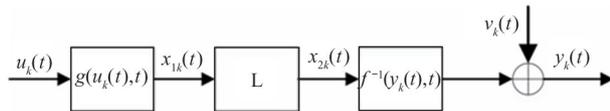


图 1 Hammerstein-Wiener 模型

$$y_k(t) = \sum_{i=1}^{n_a} a_{ik}(t) \sum_{l=1}^{n_d} d_{lk}(t) f_l^{-1}[y_k(t-i)] + \sum_{j=1}^{n_b} b_{jk}(t) \sum_{h=1}^{n_c} c_{hk}(t) g_h[u_k(t-j)] + v_k(t) \quad (1)$$

$u_k(t)$  和  $y_k(t)$  是系统  $t$  时刻的输入和输出,其中非线性函数  $f^{-1}$  和  $g$  可以用不同的函数结构来建模,如神经网络结构、多项式函数、饱和函数、死区函数等。文中把其看作是已知基函数的多项式线性组合,  $v_k(t)$  是  $t$  时刻零均值、方差为  $\sigma^2$  的白噪声序列,假设非线性函数  $f^{-1}$  和  $g$  可逆。

定义参数矢量:

$$\theta_k(t) = [a_{1k}(t)d_{1k}(t), \dots, a_{n_{dk}}(t)d_{1k}(t), \dots, a_{1k}(t)d_{n_{dk}}(t), \dots, a_{n_{dk}}(t)d_{n_{dk}}(t), b_{1k}(t)c_{1k}(t), \dots, b_{n_{bk}}(t)c_{1k}(t), \dots, b_{1k}(t)c_{n_{ck}}(t), \dots, b_{n_{bk}}(t)c_{n_{ck}}(t)]^T$$

信息矢量为

$$\phi_k(t) = [f_1[y_k(t-1)], \dots, f_1[y_k(t-n_a)], \dots, f_{n_d}[y_k(t-1)], \dots, f_{n_d}[y_k(t-n_a)], g_1[u_k(t-1)], \dots, g_1[u_k(t-n_b)], \dots, g_{n_c}[u_k(t-1)], \dots, g_{n_c}[u_k(t-n_b)]]^T$$

从而,式(1)可以获得回归形式如下:

$$y_k(t) = \phi_k^T(t) \theta_k(t) + v_k(t) \quad (2)$$

式(2)回归模型系统参数存在耦合现象,独立参数的数量大于未知参数的数量,无法辨识得到唯一的结果。为了确定唯一性,需对其系数进行归一化处理,通过辨识首先得到  $a_{ik}(t)d_{lk}(t)$  和  $b_{ik}(t)c_{lk}(t)$ ,取  $c_{ik}(t)$  的第 1 项  $c_{1k}(t) = 1$ ,  $d_{ik}(t)$  的第 1 项  $d_{1k}(t) = 1$ , 对其做除法可以得到  $a_{ik}(t)$  和  $b_{ik}(t)$ ,再对其做除法和算术平均处理。

$$c_{jk}(t) = \frac{\sum_{i=1}^{n_b} b_{ik}(t)c_{jk}(t)}{n_b}, d_{jk}(t) = \frac{\sum_{i=1}^{n_a} a_{ik}(t)d_{jk}(t)}{n_a},$$

可依次得到独立的参数变量<sup>[19-20]</sup>,从而获得唯一确定的估计参数  $\theta_{ik}$ 。

$$\theta_{ik} = [a_{1k}(t), a_{2k}(t), \dots, a_{n_{dk}}(t), b_{1k}(t), b_{2k}(t), \dots, b_{n_{bk}}(t), c_{2k}(t), \dots, c_{n_{ck}}(t), d_{2k}(t), \dots, d_{n_{dk}}(t)]^T$$

## 2 非线性时变系统参数辨识

系统在给定的作业区间上重复运行时,记录或量测由第 1 次到第  $k$  次运行时的输入输出数  $\{y_i(t), u_i(t), \phi_i(t), 0 \leq t \leq N, i = 1, 2, \dots, k\}$ ,对于固定的时刻,其中  $u_i(t)$  和  $y_i(t)$  为第  $i$  次运行的系统输入和输出,  $\phi_i(t)$  是第  $i$  次运行的信息向量,对于某一固定时刻,记

$$Y_k(t) = [y_0(t), \dots, y_k(t)]^T$$

$$\Phi_k(t) = [\phi_0(t), \dots, \phi_k(t)]^T$$

$$V_k(t) = [v_0(t), \dots, v_k(t)]^T$$

第  $k$  次重复运行后,系统所有运行特性的矩阵形式可表示为

$$Y_k(t) = \Phi_k(t) \theta_k(t) + V_k(t) \quad (3)$$

### 2.1 带遗忘因子迭代学习梯度算法

为获得系统时变参数的估计,首先介绍带遗忘

因子的迭代学习梯度算法,如算法 1 所示。引入遗忘因子  $\lambda, \lambda \in (0, 1]$ , 定义:

$$\mathbf{H}_k = \text{diag}[\lambda^{k-1}, \lambda^{k-2}, \dots, 1] \quad (4)$$

考虑如下的准则函数:

$$J_k(\hat{\boldsymbol{\theta}}_k(t), t) = \frac{1}{2} [\mathbf{Y}_k(t) - \boldsymbol{\Phi}_k(t) \hat{\boldsymbol{\theta}}_k(t)]^T \mathbf{H}_k [\mathbf{Y}_k(t) - \boldsymbol{\Phi}_k(t) \hat{\boldsymbol{\theta}}_k(t)] \quad (5)$$

基于负梯度搜索算法,可沿输出残差平方的负梯度方向修正,从而得到迭代学习梯度算法的参数更新率。

$$\hat{\boldsymbol{\theta}}_k(t) = \hat{\boldsymbol{\theta}}_{k-1}(t) + \frac{\boldsymbol{\phi}_k(t)}{r_k(t)} [y_k(t) - \boldsymbol{\phi}_k^T(t) \hat{\boldsymbol{\theta}}_{k-1}(t)] \quad (6)$$

$$r_k(t) = \lambda r_{k-1}(t) + \|\boldsymbol{\phi}_k(t)\|^2 \quad (7)$$

定义跟踪误差  $e_k(t)$  和参数误差  $\delta_k(t)$  为

$$e_k(t) = y_k(t) - \boldsymbol{\phi}_k^T(t) \hat{\boldsymbol{\theta}}_{k-1}(t) \quad (8)$$

$$\delta_k(t) = \ln \left( \frac{\|\hat{\boldsymbol{\theta}}_k(t) - \boldsymbol{\theta}_0\|}{\|\boldsymbol{\theta}_0\|} \right) \quad (9)$$

其中  $r_0(t) = 1$ , 式(6)及式(7)便构成了该模型带遗忘因子迭代学习梯度算法。当  $\lambda = 1$  时,该算法简化为一般的迭代学习梯度算法。

#### 算法 1 带遗忘因子迭代学习梯度算法

输入: 重复激励的一组数列

输出: 堆积的输出向量  $\mathbf{Y}_k(t)$

1. 对于所有的  $t = 0, 1, \dots, N$ , 给定参数估计的初始值  $\boldsymbol{\theta}_0 = 0$ , 并置  $k = 0$ ;
2. While  $k \leq K_{\max}$  ( $K_{\max}$  为最大迭代次数)
3. for each  $t \in [0, N]$
4. 在第  $k$  次重复运行时,采集输入数据  $u_k(t)$ , 计算输出数据  $y_k(t)$ , 置  $\lambda = 0.8, r_0(t) = 1$ ;
5. 通过式(6)计算得出  $\hat{\boldsymbol{\theta}}_k(t)$ ;
6. 利用式(7)更新  $r_k(t)$ ;
7. 利用式(8)和(9)计算系统跟踪误差和参数误差;
8. end
9. 检验迭代停止条件, 满足则停止; 否则置  $k = k + 1$ , 并回到步骤 3;
10. End

## 2.2 带遗忘因子迭代学习最小二乘算法

递推最小二乘算法解决了最小二乘法中数据量

较大、内存占用多等问题。同时,在此基础上引入遗忘因子增强新数据权重,防止数据饱和,对于时变系统可以获得更好的参数跟踪效果。下面介绍沿重复轴的带遗忘因子迭代学习最小二乘算法<sup>[21]</sup>,如算法 2 所示。引入遗忘因子  $\lambda, \lambda \in (0, 1]$ , 定义:

$$\mathbf{H}_k = \text{diag}[\lambda^{k-1}, \lambda^{k-2}, \dots, 1] \quad (10)$$

由式(3)可得如下准则函数:

$$J_k(\hat{\boldsymbol{\theta}}_k(t), t) = \frac{1}{2} [\mathbf{Y}_k(t) - \boldsymbol{\Phi}_k(t) \hat{\boldsymbol{\theta}}_k(t)]^T \mathbf{H}_k [\mathbf{Y}_k(t) - \boldsymbol{\Phi}_k(t) \hat{\boldsymbol{\theta}}_k(t)] \quad (11)$$

使上式最小,令  $J_k(\hat{\boldsymbol{\theta}}_k(t), t) = 0$ , 可得:

$$\hat{\boldsymbol{\theta}}_k(t) = [\boldsymbol{\Phi}_k^T(t) \mathbf{H}_k \boldsymbol{\Phi}_k(t)]^{-1} \boldsymbol{\Phi}_k^T(t) \mathbf{H}_k \mathbf{Y}_k(t) \quad (12)$$

假设  $\boldsymbol{\Phi}_k^T(t) \mathbf{H}_k \boldsymbol{\Phi}_k(t)$  可逆,定义:

$$P_k^{-1}(t) = \boldsymbol{\Phi}_k^T(t) \mathbf{H}_k \boldsymbol{\Phi}_k(t) \quad (13)$$

可得:

$$P_k^{-1}(t) = \lambda P_{k-1}^{-1}(t) + \boldsymbol{\phi}_k(t) \boldsymbol{\phi}_k^T(t) \quad (14)$$

利用矩阵求逆公式:

$$(\mathbf{A} + \mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{B} (\mathbf{C}^{-1} + \mathbf{DA}^{-1} \mathbf{B})^{-1} \mathbf{DA}^{-1}$$

取  $\mathbf{A} = P_k^{-1}(t), \mathbf{B} = \boldsymbol{\phi}_k(t), \mathbf{C} = \mathbf{I}, \mathbf{D} = \boldsymbol{\phi}_k^T(t)$ , 可得:

$$\lambda P_k(t) = P_{k-1}(t) - \frac{P_{k-1}(t) \boldsymbol{\phi}_k(t) \boldsymbol{\phi}_k^T(t) P_{k-1}(t)}{\lambda + \boldsymbol{\phi}_k^T(t) P_{k-1}(t) \boldsymbol{\phi}_k(t)} \quad (15)$$

上式两端同乘  $\boldsymbol{\phi}_k(t)$ , 除以  $\lambda$  可得:

$$P_k(t) \boldsymbol{\phi}_k(t) = \frac{P_{k-1}(t) \boldsymbol{\phi}_k(t)}{\lambda + \boldsymbol{\phi}_k^T(t) P_{k-1}(t) \boldsymbol{\phi}_k(t)} \quad (16)$$

根据式(12)和  $P_k^{-1}(t)$  的定义可得:

$$\hat{\boldsymbol{\theta}}_k(t) = P_k(t) \boldsymbol{\Phi}_k^T(t) \mathbf{H}_k \mathbf{Y}_k(t) \quad (17)$$

$$\hat{\boldsymbol{\theta}}_k(t) = P_k(t) [\lambda \boldsymbol{\Phi}_{k-1}^T(t) \mathbf{H}_{k-1} \mathbf{Y}_{k-1}(t) + \boldsymbol{\phi}_k(t) y_k(t)] \quad (18)$$

其中,  $\boldsymbol{\Phi}_{k-1}^T(t) \mathbf{H}_{k-1} \mathbf{Y}_{k-1}(t) = P_{k-1}^{-1}(t) \hat{\boldsymbol{\theta}}_{k-1}(t)$

$$\hat{\boldsymbol{\theta}}_k(t) = P_k(t) [\lambda P_{k-1}^{-1}(t) \hat{\boldsymbol{\theta}}_{k-1}(t) + \boldsymbol{\phi}_k(t) y_k(t)] \quad (19)$$

$$\hat{\boldsymbol{\theta}}_k(t) = P_k(t) [(P_k^{-1}(t) - \boldsymbol{\phi}_k(t) \boldsymbol{\phi}_k^T(t)) \hat{\boldsymbol{\theta}}_{k-1}(t) + \boldsymbol{\phi}_k(t) y_k(t)] \quad (20)$$

$$\hat{\boldsymbol{\theta}}_k(t) = \hat{\boldsymbol{\theta}}_{k-1}(t) - P_k(t) \boldsymbol{\phi}_k(t) \boldsymbol{\phi}_k^T(t) \hat{\boldsymbol{\theta}}_{k-1}(t)$$

$$+ P_k(t) \boldsymbol{\phi}_k(t) y_k(t) \quad (21)$$

$$\hat{\boldsymbol{\theta}}_k(t) = \hat{\boldsymbol{\theta}}_{k-1}(t) + P_k(t) \boldsymbol{\phi}_k(t) e_k(t) \quad (22)$$

根据式(20)和式(22)可得:

$$\hat{\boldsymbol{\theta}}_k(t) = \hat{\boldsymbol{\theta}}_{k-1}(t) + \frac{P_{k-1}(t) \boldsymbol{\phi}_k(t)}{\lambda + \boldsymbol{\phi}_k^T(t) P_{k-1}(t) \boldsymbol{\phi}_k(t)} e_k(t) \quad (23)$$

定义跟踪误差  $e_k(t)$  和参数误差  $\delta_k(t)$ :

$$e_k(t) = y_k(t) - \boldsymbol{\phi}_k(t) \hat{\boldsymbol{\theta}}_{k-1}(t) \quad (24)$$

$$\delta_k(t) = \ln \left( \frac{\|\hat{\boldsymbol{\theta}}_k(t) - \boldsymbol{\theta}\|}{\|\boldsymbol{\theta}\|} \right) \quad (25)$$

式(15)及式(23)构成该模型带遗忘因子迭代学习最小二乘算法。其中,当  $\lambda = 1$  时,该算法简化为普通的迭代学习最小二乘算法。

---

**算法 2** 带遗忘因子迭代学习最小二乘算法

---

输入: 重复激励的一组数列

输出: 堆积的输出向量  $\mathbf{Y}_k(t)$

---

1. 对于所有的  $t = 0, 1, \dots, N$ , 给定参数估计的初始值  $\hat{\boldsymbol{\theta}}_0 = \mathbf{0}$ , 并置  $k = 0$ ;
  2. While  $k \leq K_{\max}$  ( $K_{\max}$  为最大迭代次数)
  3. for each  $t \in [0, N]$
  4. 在第  $k$  次重复运行时, 采集输入数据  $u_k(t)$ , 计算输出数据  $y_k(t)$ , 并置  $P_k(t) = 10^6 I$ ;
  5. 利用式(15)更新  $P_k(t)$ ;
  6. 通过式(23)计算得出  $\hat{\boldsymbol{\theta}}_k(t)$ ;
  7. 利用式(24)和(25)计算系统跟踪误差和参数误差;
  8. end
  9. 检验迭代停止条件, 满足则停止; 否则置  $k = k + 1$ , 并回到步骤 3;
  10. End
- 

### 3 算例仿真

本节将完成具体的算例仿真, 考虑如下时变 Hammerstein 模型:

$$y_k(t) = \sum_{i=1}^{n_a} a_{ik}(t) \sum_{l=1}^{n_d} d_{lk}(t) f_l^{-1}[y_k(t-i)] + \sum_{j=1}^{n_b} b_{jk}(t) \sum_{h=1}^{n_c} c_{hk}(t) g_h[u_k(t-j)] + v_k(t)$$

分别采用带遗忘因子迭代学习梯度算法和带遗忘因子迭代学习最小二乘算法对该模型进行参数估计, 仿真中数据长度  $t = 1000$ , 迭代次数取  $k = 500$ 。

为了说明本文所提学习算法的有效性, 引入现有带遗忘因子梯度算法和带遗忘因子递推最小二乘算法对该模型进行参数估计, 仿真中数据长度  $t = 500$ 。仿真时,  $v_k(t)$  采用均值为零、方差  $\sigma^2 = 0.1^2$  的随机白噪声序列, 对应的信噪比为 7.66%, 输入  $u_k(t)$  采用零均值、不相关可测随机信号序列。输入输出部分非线性函数取  $f_1^{-1}[y_k(t)] = y_k(t)$ ,  $g_1[u_k(t)] = u_k(t)$ ,  $f_2^{-1}[y_k(t)] = \sin(y_k(t) + 1)$ ,  $g_2[u_k(t)] = u_k^2(t)$ 。为避免参数耦合问题, 线性部分参数:

$$\begin{aligned} a_{1k}(t) &= (1.5^{\sin(\frac{1000}{t})} - 1) \\ &\quad \times (0.4 + 0.029 \times \sin(0.022 \times k)) \\ a_{2k}(t) &= |0.1 \times \sin(0.05 \times t)| \\ &\quad + (0.3 + 0.031 \times \sin(0.011 \times k)) \\ b_{1k}(t) &= -1 + 0.4 \times \sin\left(\frac{\pi}{\|0.9 \times t - 60.5\|}\right) \\ &\quad + 0.012 \times \sin(0.034 \times k) \\ b_{2k}(t) &= (\log(0.1 + 1.4 \times \|\cos\left(\frac{10 \times \pi}{t}\right)\|)) \\ &\quad + 0.1) \times (0.7 + 0.054 \times \sin(0.028 \times k)) \end{aligned}$$

非线性部分:

$$\begin{aligned} c_{1k}(t) &= 1, d_{1k}(t) = 1 \\ c_{2k}(t) &= (0.23 \times \sin(0.032 \times t)) \\ &\quad + 0.042 \times \sin(0.002 \times k) \\ c_{3k}(t) &= -1.5 + 0.5 \times (1.5 \times t + 130)^{\frac{1}{5}} \\ &\quad + 0.013 \times \sin(0.005 \times k) \\ d_{2k}(t) &= (0.1 + 0.2 \times \cos(\pi \times 0.07 \times t)) \\ &\quad + 0.045 \times \cos(0.003 \times k) \end{aligned}$$

取  $\lambda = 0.8$ , 初始化  $\boldsymbol{\theta}_0 = [0, 0, 0, 0, 0, 0, 0]^T$ ,  $P_0(t) = I \times 10^6$ ,  $r_0(t) = 1$ 。对于固定的第  $k$  次迭代, 选取在所有时刻  $t$  上, 跟踪误差绝对值最大的时刻  $t_1$  和参数误差绝对值最大的时刻  $t_2$ ,  $t_1$  和  $t_2$  时刻对应的值  $e_k(t_1)$  和  $\delta_k(t_2)$  分别作为该次迭代的跟踪误差和参数误差。定义参数误差  $\delta_k(t)$  和跟踪误差  $e_k(t)$ :

$$\delta_k(t) = \ln \max_{0 \leq t \leq N} \left( \frac{\|\hat{\boldsymbol{\theta}}_k(t) - \boldsymbol{\theta}_0\|}{\|\boldsymbol{\theta}_0\|} \right)$$

$$e_k(t) = \ln \max_{0 \leq t \leq N} (y_k(t) - \boldsymbol{\varphi}_k^T(t) \hat{\boldsymbol{\theta}}_k(t-1))$$

由于三维曲面上难以直观表达采用迭代学习算法所获得的参数估计情况, 以迭代次数  $k = 500$ 、时

间轴  $t = 100$  时刻前的数据, 作为沿时间轴快变参数估计结果, 以及  $t = 1000$ 、迭代次数  $k = 100$  前的数据, 作为沿迭代轴缓变参数估计的结果来验证算法的性能。通过估计得到的参数, 分析算法对该模型时变参数沿时间轴和迭代轴的跟踪效果。如图 2 ~ 7 所示。

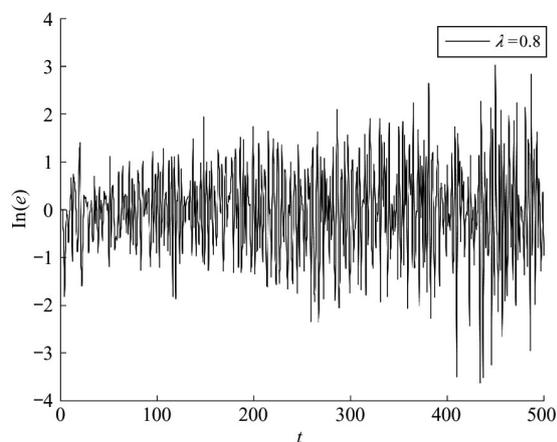


图 2 带遗忘因子梯度算法跟踪误差

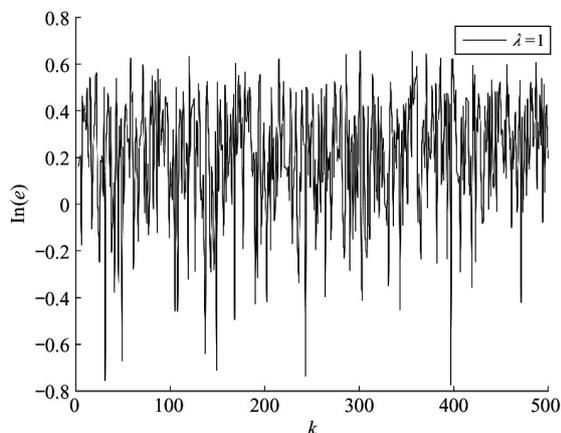


图 3 迭代学习梯度算法跟踪误差

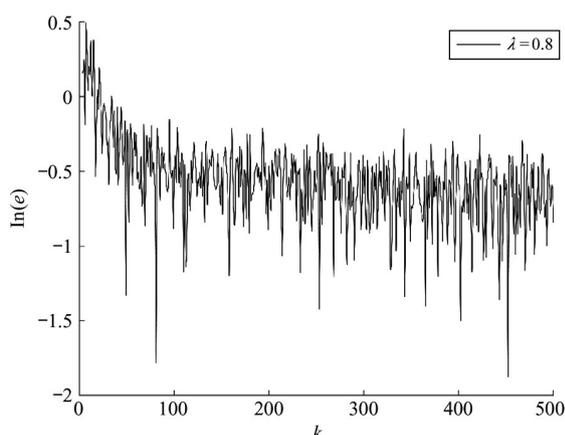


图 4 带遗忘因子迭代学习梯度算法跟踪误差

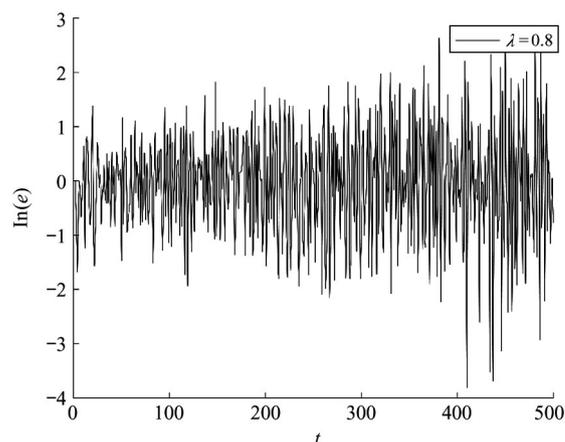


图 5 带遗忘因子递推最小二乘算法跟踪误差

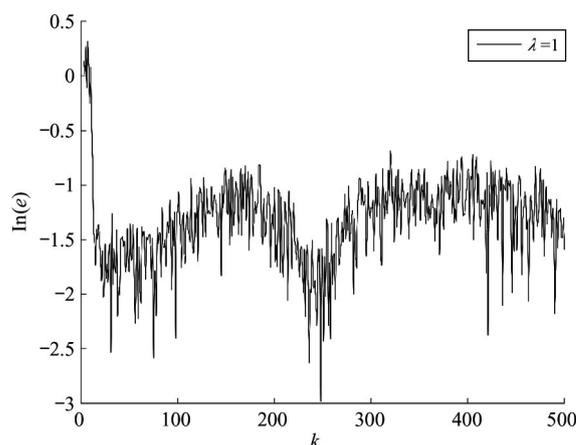


图 6 迭代学习最小二乘算法跟踪误差

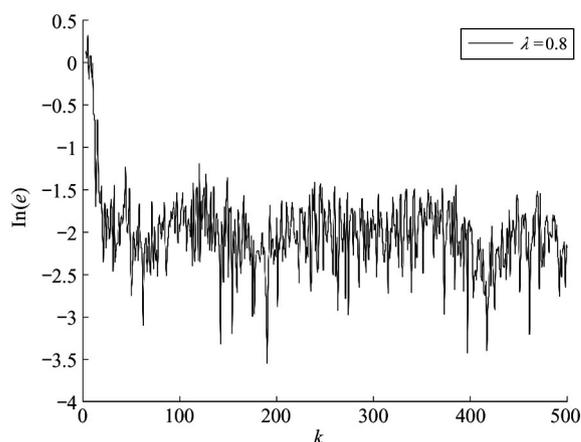


图 7 带遗忘因子迭代学习最小二乘算法跟踪误差

仿真结果显示, 带遗忘因子梯度算法和带遗忘因子递推最小二乘算法无法实现对该模型时变参数的完全估计; 而迭代学习梯度算法和迭代学习最小二乘算法均可较好地实现对该模型时变参数的跟踪。引入遗忘因子进一步改善了迭代轴时变参数的

辨识效果。如图8~15所示。

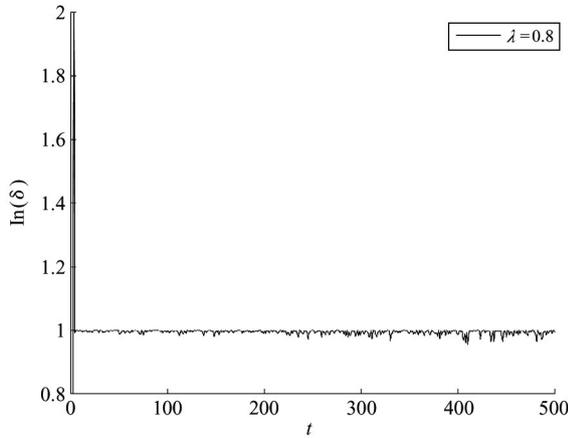


图8 带遗忘因子梯度算法参数误差

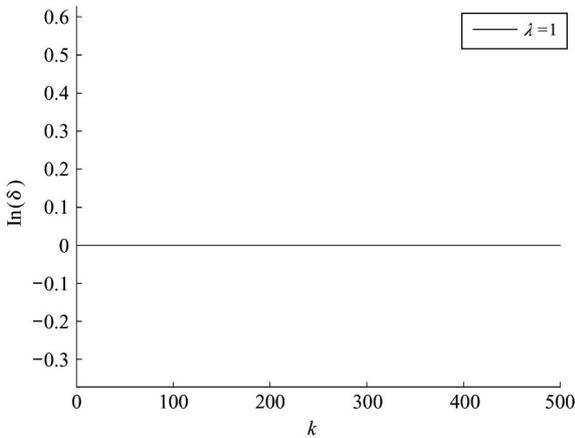


图9 迭代学习梯度算法参数误差

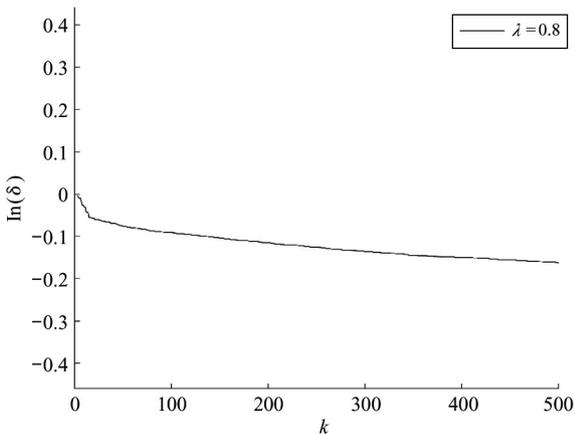


图10 带遗忘因子迭代学习梯度算法参数误差

#### 4 结论

针对时变 Hammerstein-Wiener 模型的辨识,本研究采用了沿迭代轴的重复学习算法,该算法可用

于解决时变系统的参数估计问题。针对文中给定的时变 Hammerstein-Wiener 模型,假设输出非线性部分可逆,通过算例仿真证明系统在重复持续激励的条件下,传统的沿时间轴递推算法估计该模型时变

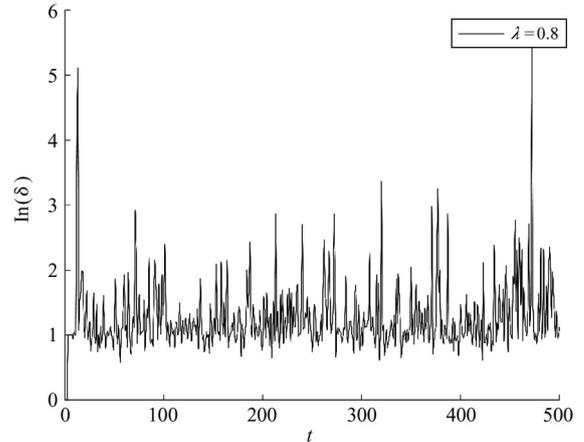


图11 带遗忘因子递推最小二乘算法参数误差

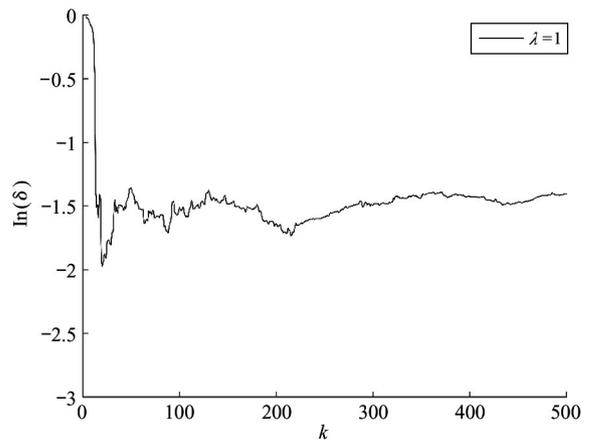


图12 迭代学习最小二乘算法参数误差

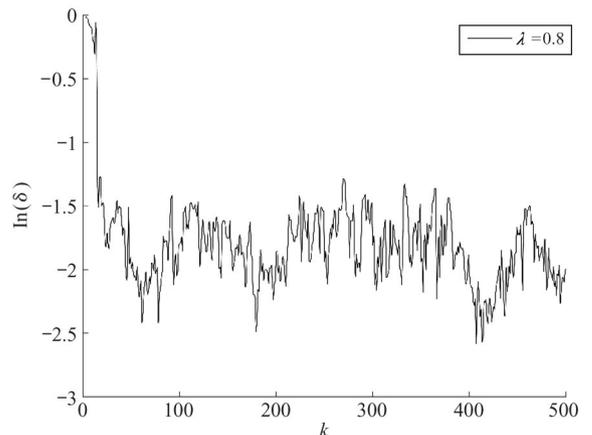


图13 带遗忘因子迭代学习最小二乘算法参数误差

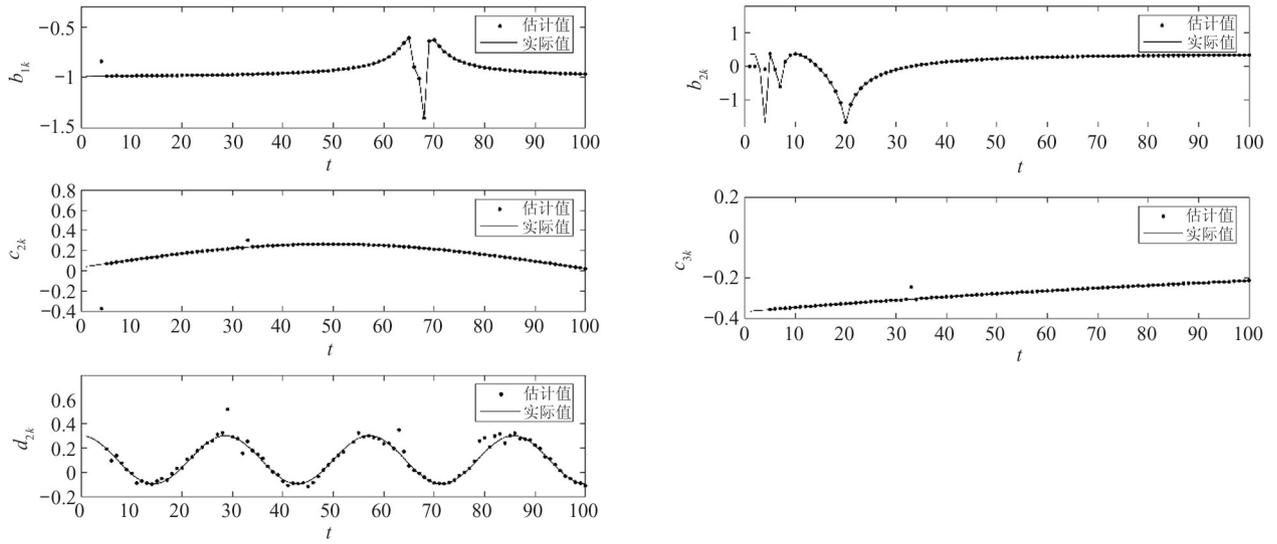


图 14 采用带遗忘因子迭代学习最小二乘法估计沿时间轴快变参数的结果

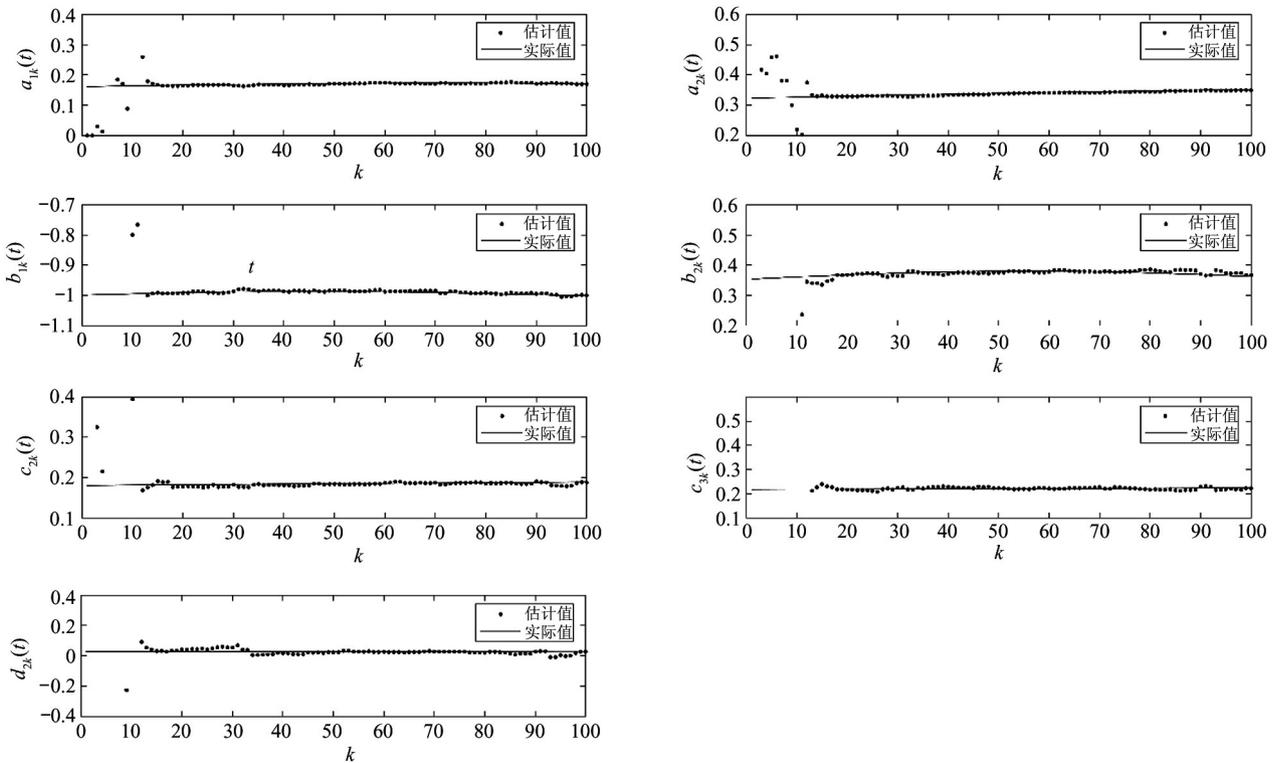


图 15 采用带遗忘因子迭代学习最小二乘法估计沿迭代轴缓变参数的结果

参数时误差不收敛,无法获得参数的完全估计;迭代学习梯度算法和迭代学习最小二乘法可以实现对时变参数的估计,同时引入遗忘因子可以有效解决沿迭代轴发生缓变参数的估计问题。

参考文献

[ 1 ] BAI E W. An optimal two-stage identification algorithm for Hammerstein-Wiener nonlinear systems[J]. Automat-

ica, 1998,34(3):333-338.

[ 2 ] BAIE W. A blind approach to the Hammerstein-Wiener model identification[J]. Automatica, 2002,38(6):967-979.

[ 3 ] GIRI F, BAI E W. Block-oriented nonlinear system identification[M]. London: Springer, 2010;3-11.

[ 4 ] 丁锋. 系统辨识新论[M]. 北京:科学出版社, 2013: 37-38.

- [ 5 ] MEHMOOD A, CHAUDHARY N I, ZAMEER A, et al. Backtracking search optimization heuristics for nonlinear Hammerstein controlled auto regressive auto regressive systems[J]. ISA Transactions, 2019,91:99-113.
- [ 6 ] AWRYNCZUK M, TATJEWSKI P. Offset-free state-space nonlinear predictive control for wiener systems[J]. Information Sciences, 2020,511:127-151.
- [ 7 ] RAJA M, SHAH F H, ALAIDAROUS E S, et al. Design of bio-inspired heuristic technique integrated with interior-point algorithm to analyze the dynamics of heartbeat model[J]. Applied Soft Computing, 2016,52:605-629.
- [ 8 ] ZHU Y. Estimation of an NLN Hammerstein-Wiener model [J]. Automatica, 2002,38(9):1607-1614.
- [ 9 ] 李妍, 毛志忠, 王琰, 等. 基于偏差补偿递推最小二乘的 Hammerstein-Wiener 模型辨识[J]. 自动化学报, 2010,36(1):163-168.
- [ 10 ] DU N, ZHANG L, LONG X, et al. Recursive identification for choke finger system in wind tunnel [J]. ISA Transactions, 2020,107(2):173-180.
- [ 11 ] DING F, WANG F, XU L, et al. Decomposition based least squares iterative identification algorithm for multivariate pseudo-linear ARMA systems using the data filtering [J]. Journal of the Franklin Institute, 2016,354(3):1321-1339.
- [ 12 ] LIJ H, ZONG T C, LU G P. Parameter identification of Hammerstein-Wiener nonlinear systems with unknown time delay based on the linear variable weight particle swarm optimization[J]. ISA Transactions, 2021,120:89-98.
- [ 13 ] BROURI A. Identification of nonlinear system composed of parallel coupling of wiener and Hammerstein models [J]. Asian Journal of Control, 2021,1:1-13.
- [ 14 ] MZYK G, WACHEL P. Kernel-based identification of Wiener-Hammerstein system[J]. Automatica, 2017,83:275-281.
- [ 15 ] RISULEO R S, HJALMARSSON H. Nonparametric models for Hammerstein-Wiener and Wiener-Hammerstein system identification [J]. IFAC-PapersOnLine, 2020,53(2):400-405.
- [ 16 ] 孙明轩, 毕宏博. 学习辨识: 最小二乘算法及其重复一致性[J]. 自动化学报, 2012,38(5):698-706.
- [ 17 ] 孙明轩, 毕宏博. 最小二乘学习辨识[C]//第30届中国控制会议, 烟台:中国自动化学会控制理论专业委员会 D 卷, 2011:969-974.
- [ 18 ] 仲国民, 孙明轩. 一类 Wiener 非线性时变系统的迭代学习辨识[J]. 电子与信息学报, 2021,43(9):2594-2600.
- [ 19 ] DING F, LIU X P, LIU G J. Identification methods for Hammerstein nonlinear systems [J]. Digital Signal Processing, 2011,21(2):215-238.
- [ 20 ] YU F, MAO Z Z, YUAN P, et al. Recursive parameter estimation for Hammerstein-Wiener systems using modified EKF algorithm [J]. ISA Transactions, 2017,70:104-115.
- [ 21 ] 萧德云. 系统辨识理论及应用[M]. 北京:清华大学出版社, 2014:120-124.

## Iterative learning algorithm with forgetting factor for Hammerstein-Wiener time-varying systems

ZHONG Guomin, YU Qile, WANG Liming

(College of Information Engineering, Zhejiang University of Technology, Hangzhou 310023)

### Abstract

For Hammerstein-Wiener nonlinear time-varying systems running repeatedly on a finite interval, the nonlinear output part of Hammerstein-Wiener system is tackled based on polynomial expansion to construct the regression model. The time-varying parameters of the system are estimated by the iterative learning gradient algorithm with forgetting factor and the iterative learning least square algorithm with forgetting factor. When the system parameters change rapidly along the time axis and slowly along the iteration axis, the forgetting factor is modified to improve the identification accuracy. The derivation processes of the two algorithms are given and verified by numeral simulations. The simulation results show that the iterative learning least squares algorithm with forgetting factor can perform high identification accuracy, fast convergence speed and few iterations when reaching the same performance index, which verifies the effectiveness of the proposed learning algorithm.

**Key words:** learning identification, least square, stochastic gradient, Hammerstein-Wiener model