第 30 卷 第 8 期
Vol. 30 No. 8

控 制 与 决 策
Control and Decision

2015 年 8 月
Aug. 2015

文章编号: 1001-0920 (2015) 08-1491-06 **DOI:** 10.13195/j.kzyjc.2014.0973

非均匀 Hammerstein-Wiener 系统的递阶随机梯度辨识算法

刘冉冉[1,2], 潘天红[1], 李正明[1]

(1. 江苏大学 电气信息工程学院，江苏 镇江 212013；2. 江苏理工学院 汽车与交通工程学院，江苏 常州 213001)

摘 要: 针对一类非均匀数据采样 Hammerstein-Wiener 系统, 提出一种递阶多新息随机梯度算法. 首先基于提升技术, 推导出系统的状态空间模型, 并考虑因果约束关系, 将该模型分解成两个子系统, 利用多新息遗忘随机梯度算法辨识出此模型的参数; 然后, 引入可变遗忘因子, 提出一种修正函数并在线确定其大小, 提高了算法的收敛速度及抗干扰能力. 仿真实例验证了所提出算法的有效性和优越性.

关键词: 参数估计；随机梯度；Hammerstein-Wiener 系统；递阶辨识；可变遗忘因子

中图分类号: TP273 **文献标志码:** A

Hierarchical stochastic gradient identification for Hammerstein-Wiener systems with non-uniformly sampling

LIU Ran-ran[1,2], *PAN Tian-hong*[1], *LI Zheng-ming*[1]

(1. School of Electrical and Information Engineering，Jiangsu University，Zhenjiang 212013，China；2. School of Automotive and Traffic Engineering，Jiangsu University of Technology，Changzhou 213000，China．Correspondent: PAN Tian-hong，E-mail：thpan@ujs.edu.cn)

Abstract: A hierarchical multi-innovation stochastic gradient identification algorithm is proposed for Hammerstein-Wiener(H-W) nonlinear systems with non-uniformly sampling. The corresponding state space models of H-W are derived by using the lifting technique. Considering the causality constraints, the H-W system is decomposed into two subsystems firstly. Then the model parameters are identified by using the multi-innovation based stochastic gradient algorithm with forgetting factors. In order to improve the convergent rate and the disturbance rejection, a new kind of variable forgetting factor algorithm is also presented. Simulation examples demonstrate that the proposed algorithm has fast convergence speed and is robust to the noise.

Keywords: parameters estimation；stochastic gradient；Hammerstein-Wiener system；hierarchical identification；varying forgetting factor

0 引 言

在工业生产过程中, 受传感器采样频率、物理设备的特性, 以及机械、环境因素等影响, 使得系统采样呈现不规则状态, 即多率系统. 由于该系统能够反映实际工业生产中数据采样特点, 且直接影响控制系统性能, 受到了工业界与学术界的普遍关注. 对于线性非均匀多率系统而言, 利用提升技术, 结合多新息递阶辨识原理与最小二乘算法, 在不同噪声条件下, 基于不同的模型结构 (如: CAR、CARIMA、Box-Jenkins 等), 涌现出一系列辨识方法, 并取得了良好的辨识效果[1-7].

然而, 实际的工业过程几乎都呈现不同的非线性, 很难直接使用上述算法. 文献 [8] 基于多项式转换技术, 提出了一种辨识 Hammerstein 非线性双率系统的最小二乘辨识算法, 但提升后的系统参数向量维数大, 且在参数分离过程中大量参数被重复估计, 导致计算量较大. 文献 [9-10] 针对具有预负载特性的双率非线性系统, 通过关键项分离将系统模型转换为辨识模型, 并利用多项式转换技术, 提出两种改进的随机梯度算法. 但上述算法不仅待辨识参数维数高, 而且在有色噪声干扰下, 无法直接由辨识参数计算出原系统参数.

收稿日期: 2014-06-18；**修回日期:** 2014-08-30.

基金项目: 国家自然科学基金项目(61273142)；江苏省自然科学基金项目(BK2011466)；江苏省研究生培养创新工程项目(CXLX12_0648)；江苏省六大人才高峰项目(2012-DZXX-045)；江苏省高校优势学科建设工程项目(PAPD).

作者简介: 刘冉冉(1982−), 女, 博士生, 从事非线性多率系统辨识的研究；潘天红(1974−), 男, 教授, 博士生导师, 从事非线性系统模型辨识与优化控制等研究.

为此, 本文基于文献 [11-12] 的方法, 以非均匀采样数据 Hammerstein-Wiener (H-W) 非线性系统为对象, 提出一种带可变遗忘因子的递阶多新息随机梯度算法. 该方法利用递阶辨识原理, 将系统分解为两个子系统, 直接对系统参数进行辨识, 从而降低计算负荷; 同时引入可变遗忘因子, 以提高算法的抗干扰能力及辨识精度.

1 问题描述

一类非均匀采样数据 H-W 非线性系统结构如图 1 所示. 其中: $f(\cdot)$、$h(\cdot)$ 分别为输入和输出静态非线性模块, P_c 为动态线性模块.

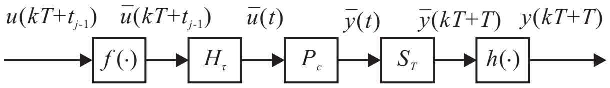


图 1 非均匀采样 H-W 系统

设静态非线性模块 $h(\cdot)$ 可逆, 则有

$$\bar{u}(kT+t_{j-1})=f(u(kT+t_{j-1}))=$$
$$c_1f_1(u(kT+t_{j-1}))+\cdots+c_{n_c}f_{n_c}(u(kT+t_{j-1}))=$$
$$\sum_{s=1}^{n_c}c_sf_s(u(kT+t_{j-1})),\tag{1}$$
$$\bar{y}(kT)=h^{-1}(y(kT))=g(y(kT))=$$
$$d_1g_1(y(kT))+\cdots+d_{n_d}g_{n_d}(y(kT))=$$
$$\sum_{l=1}^{n_d}d_lg_l(y(kT)).\tag{2}$$

其中: $f_1(\cdot),\cdots,f_{n_c}(\cdot)$ 和 $g_1(\cdot),\cdots,g_{n_d}(\cdot)$ 为已知基函数; $c_1,\cdots,c_{n_c}$ 和 $d_1,\cdots,d_{n_d}$ 为静态非线性模块的未知参数;

$$\bar{u}(t)=$$
$$\begin{cases}\bar{u}(kT),\ kT\leqslant t<kT+t_1;\\ \bar{u}(kT+t_1),\ kT+t_1\leqslant t<kT+t_2;\\ \quad\vdots\\ \bar{u}(kT+t_{m-1}),\ kT+t_{m-1}\leqslant t<kT+T;\end{cases}=$$
$$\begin{cases}\sum\limits_{s=1}^{n_c}c_sf_s(u(kT)),\ kT\leqslant t<kT+t_1;\\ \sum\limits_{s=1}^{n_c}c_sf_s(u(kT+t_1)),\ kT+t_1\leqslant t<kT+t_2;\\ \quad\vdots\\ \sum\limits_{s=1}^{n_c}c_sf_s(u(kT+t_{m-1})),\ kT+t_{m-1}\leqslant t<kT+T.\end{cases}\tag{3}$$

在周期 $[kT,kT+T]$ 上, 非均匀零阶保持器 H_τ 控制输入 $\bar{u}$ 非均匀刷新 m 次, 即控制系统的输入信号 u 非均匀刷新 m 次, 时间间隔为 $\{\tau_1,\tau_2,\cdots,\tau_m\}$, 则输入刷新时刻点依次为 $kT+t_{j-1},j=1,2,\cdots,m$, 其中 $t_{j-1}:=\tau_1+\tau_2+\cdots+\tau_{j-1}$. 为方便起见, 令 $t_0=0$, 则系统框架周期为 $T=t_m=\tau_1+\tau_2+\cdots+\tau_m$, 经提升技术处理后特性如式 (3) 所示.

设 P_c 状态空间模型为

$$P_c:=\begin{cases}\dot{\boldsymbol{x}}(t)=\boldsymbol{A}_1\boldsymbol{x}(t)+\boldsymbol{B}_1\bar{u}(t),\\ \bar{y}(t)=\boldsymbol{C}x(t)+D\bar{u}(t).\end{cases}\tag{4}$$

其中: $\boldsymbol{x}(t)\in R^n$ 为状态向量; $y(t)\in R$, $\bar{u}(t)\in R$ 分别为系统的输出和输入; $\boldsymbol{A}_1\in R^{n\times n}$, $\boldsymbol{B}_1\in R^n$, $\boldsymbol{C}\in R^{1\times n}$, $D\in R$ 分别为系统的参数矩阵. 非均匀刷新的输入数据为 $u(kT+t_{j-1}),j=1,2,\cdots,m$, 周期采样的输出数据为 $y(kT)$, 以周期 T 离散化系统 (4)[13], 得

$$y(kT)=\alpha(z)\bar{y}(kT)+\sum_{j=1}^{m}\beta_j(z)\bar{u}(kT+t_{j-1}).\tag{5}$$

其中

$$\alpha(z)=\alpha_1z^{-1}+\alpha_2z^{-2}+\cdots+\alpha_{n_a}z^{-n_a},$$
$$\beta_1(z)=\beta_{10}+\beta_{11}z^{-1}+\cdots+\beta_{1n_b}z^{-n_b},$$
$$\beta_j(z)=\beta_{j1}z^{-1}+\beta_{j2}z^{-2}+\cdots+\beta_{jn_b}z^{-n_b}.$$

这里: $j=2,3,\cdots,m$; z^{-1} 为后移因子, 即 $z^{-1}x(kT)=x(kT-T)$. 不失一般性, 引入白噪声, 综合式 (1)、(2)、(5), 可得

$$y(kT)=\alpha(z)\sum_{l=1}^{n_d}d_lg_l(y(kT))+\sum_{j=1}^{m}\beta_j(z)\times$$
$$\sum_{s=1}^{n_c}c_sf_s(u(kT+t_{j-1}))+v(kT).\tag{6}$$

其中: n_a、n_b、n_c、n_d 为给定的模型阶次; 当 $k\leqslant 0$ 时, $u(kT+t_{j-1})=0$, $y(kT)=0$, $v(kT)=0$. 定义系统线性部分参数向量 $\boldsymbol{a}$、$\boldsymbol{b}$ 及非线性部分参数向量 $\boldsymbol{c}$、$\boldsymbol{d}$ 分别为

$$\boldsymbol{a}:=[\alpha_1,\alpha_2,\cdots,\alpha_{n_a}]^{\mathrm{T}}\in R^{n_a},$$
$$\boldsymbol{b}:=[\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_m]^{\mathrm{T}}\in R^{mn_b+1},$$
$$\boldsymbol{c}:=[c_1,c_2,\cdots,c_{n_c}]^{\mathrm{T}}\in R^{n_c},$$
$$\boldsymbol{d}:=[d_1,d_2,\cdots,d_{n_d}]^{\mathrm{T}}\in R^{n_d}.$$

其中

$$\boldsymbol{\beta}_1:=[\beta_{10},\beta_{11},\cdots,\beta_{1n_b}]^{\mathrm{T}}\in R^{n_b+1},$$
$$\boldsymbol{\beta}_j:=[\beta_{j1},\beta_{j2},\cdots,\beta_{jn_b}]^{\mathrm{T}}\in R^{n_b},\ j=2,3,\cdots,m.$$

于是系统的参数向量为 $\boldsymbol{\Theta}=[\boldsymbol{a},\boldsymbol{b},\boldsymbol{c},\boldsymbol{d}]^{\mathrm{T}}$.

2 带可变遗忘因子的多新息随机梯度算法

定义系统的输出与输入信息 $\boldsymbol{G}(kT)$ 和 $\boldsymbol{F}(kT)$ 分别为

$$\boldsymbol{G}(kT)=\begin{bmatrix}g_1(y(kT-T)) & \cdots & g_{n_d}(y(kT-T))\\ \vdots & \ddots & \vdots\\ g_1(y(kT-n_aT)) & \cdots & g_{n_d}(y(kT-n_aT))\end{bmatrix},\tag{7}$$

$$\boldsymbol{F}(kT)=\begin{bmatrix} f_1(u(kT)) & \cdots & f_{n_c}(u(kT) \\ \vdots & \ddots & \vdots \\ f_1(u(kT-n_bT)) & \cdots & f_{n_c}(u(kT-n_bT)) \\ f_1(u(kT-T+t_1)) & \cdots & f_{n_c}(u(kT-T+t_1)) \\ \vdots & \ddots & \vdots \\ f_1(u(kT-n_bT+t_1)) & \cdots & f_{n_c}(u(kT-n_bT+t_1)) \\ f_1(u(kT-T+t_{m-1})) & \cdots & f_{n_c}(u(kT-T+t_{m-1})) \\ \vdots & \ddots & \vdots \\ f_1(u(kT-n_bT+t_{m-1})) & \cdots & f_{n_c}(u(kT-n_bT+t_{m-1})) \end{bmatrix}. \tag{8}$$

其中: $\boldsymbol{F}(kT)\in R^{(mn_b+1)\times n_c}$, $\boldsymbol{G}(kT)\in R^{n_a\times n_d}$. 则式 (6) 可写成

$$y(kT)=\boldsymbol{a}^{\rm T}\boldsymbol{G}(kT)\boldsymbol{d}+\boldsymbol{b}^{\rm T}\boldsymbol{F}(kT)\boldsymbol{c}+v(kT). \tag{9}$$

基于递阶辨识原理, 可将非均匀采样数据 H-W 非线性系统 (9) 分解为两个子系统: 一个包含线性动态模块的参数 $\boldsymbol{\eta}:=[\boldsymbol{a},\boldsymbol{b}]^{\rm T}$; 另一个包含非线性静态模块的参数 $\boldsymbol{\gamma}:=[\boldsymbol{d},\boldsymbol{c}]^{\rm T}$. 由于参数 $\boldsymbol{a}$、$\boldsymbol{b}$、$\boldsymbol{c}$、$\boldsymbol{d}$ 未知, 在 kT 时刻, 用前一周期 $(k-1)T$ 时的估计值 $\hat{\boldsymbol{a}}_{k-1}$、$\hat{\boldsymbol{b}}_{k-1}$、$\hat{\boldsymbol{c}}_{k-1}$、$\hat{\boldsymbol{d}}_{k-1}$ 代替, 定义 $\hat{\boldsymbol{\eta}}(kT):=[\hat{\boldsymbol{a}}_k,\hat{\boldsymbol{b}}_k]^{\rm T}$, $\hat{\boldsymbol{\gamma}}(kT):=[\hat{\boldsymbol{c}}_k,\hat{\boldsymbol{d}}_k]^{\rm T}$ 分别表示 kT 时刻 $\boldsymbol{\eta}$、$\boldsymbol{\gamma}$ 的估计值, 定义广义信息向量 $\varphi(\hat{\boldsymbol{c}}_{k-1}\hat{\boldsymbol{d}}_{k-1},kT)\in R^{n_a+mn_b+1}$、$\psi(\hat{\boldsymbol{a}}_{k-1}\hat{\boldsymbol{b}}_{k-1},kT)\in R^{n_c+n_d}$ 分别为

$$\varphi(\hat{\boldsymbol{c}}_{k-1}\hat{\boldsymbol{d}}_{k-1},kT)=\begin{bmatrix}\boldsymbol{G}(kT)\hat{\boldsymbol{d}}_{k-1}\\ \boldsymbol{F}(kT)\hat{\boldsymbol{c}}_{k-1}\end{bmatrix}, \tag{10}$$

$$\psi(\hat{\boldsymbol{a}}_{k-1}\hat{\boldsymbol{b}}_{k-1},kT)=\begin{bmatrix}\boldsymbol{G}^{\rm T}(kT)\hat{\boldsymbol{a}}_{k-1}\\ \boldsymbol{F}^{\rm T}(kT)\hat{\boldsymbol{b}}_{k-1}\end{bmatrix}. \tag{11}$$

于是系统 (9) 可以分解为如下两个子系统:

$$M_1:\ \ y(kT)=\varphi^{\rm T}(\hat{\boldsymbol{c}}_{k-1}\hat{\boldsymbol{d}}_{k-1},kT)\boldsymbol{\eta}+v(kT), \tag{12}$$

$$M_2:\ \ y(kT)=\psi^{\rm T}(\hat{\boldsymbol{a}}_{k-1}\hat{\boldsymbol{b}}_{k-1},kT)\boldsymbol{\gamma}+v(kT). \tag{13}$$

M_1 中只含线性动态参数 $\boldsymbol{\eta}$, M_2 中只含非线性静态参数 $\boldsymbol{\gamma}$, 两个子系统通过向量 $\varphi(\hat{\boldsymbol{c}}_{k-1}\hat{\boldsymbol{d}}_{k-1},kT)$、$\psi(\hat{\boldsymbol{a}}_{k-1}\hat{\boldsymbol{b}}_{k-1},kT)$ 相关联. 为了充分利用系统历史采样数据, 基于多新息理论, 定义信息矩阵 $\boldsymbol{\Phi}(\hat{\boldsymbol{c}}_{k-1}\hat{\boldsymbol{d}}_{k-1},kT)\in R^{p\times(n_a+mn_b+1)}$、$\boldsymbol{\Psi}(\hat{\boldsymbol{a}}_{k-1}\hat{\boldsymbol{b}}_{k-1},kT)\in R^{p\times(n_c+n_d)}$ 以及堆积输出向量 $\boldsymbol{Y}(p,kT)\in R^p$ 分别为

$$\boldsymbol{Y}(p,kT)=\begin{bmatrix}y(kT)\\ y(kT-T)\\ \vdots\\ y(kT-(p-1)T)\end{bmatrix}, \tag{14}$$

$$\boldsymbol{\Phi}(\hat{\boldsymbol{c}}_{k-1}\hat{\boldsymbol{d}}_{k-1},kT)=\begin{bmatrix}\varphi(\hat{\boldsymbol{c}}_{k-1}\hat{\boldsymbol{d}}_{k-1},kT)\\ \vdots\\ \varphi(\hat{\boldsymbol{c}}_{k-1}\hat{\boldsymbol{d}}_{k-1},kT-(p-1)T)\end{bmatrix}, \tag{15}$$

$$\boldsymbol{\Psi}(\hat{\boldsymbol{a}}_{k-1}\hat{\boldsymbol{b}}_{k-1},kT)=\begin{bmatrix}\psi(\hat{\boldsymbol{a}}_{k-1}\hat{\boldsymbol{b}}_{k-1},kT)\\ \vdots\\ \psi(\hat{\boldsymbol{a}}_{k-1}\hat{\boldsymbol{b}}_{k-1},kT-(p-1)T)\end{bmatrix}, \tag{16}$$

其中 p 为新息长度.

以辨识误差平方和为准则函数, 即

$$J_1(\boldsymbol{\eta})=||\boldsymbol{Y}(p,kT)-\boldsymbol{\Phi}^{\rm T}(\hat{\boldsymbol{c}}_{k-1}\hat{\boldsymbol{d}}_{k-1},kT)\boldsymbol{\eta}||^2, \tag{17}$$

$$J_2(\boldsymbol{\gamma})=||\boldsymbol{Y}(p,kT)-\boldsymbol{\Psi}^{\rm T}(\hat{\boldsymbol{a}}_{k-1}\hat{\boldsymbol{b}}_{k-1},kT)\boldsymbol{\gamma}||^2, \tag{18}$$

其中 $||\cdot||^2$ 为向量范数.

J_1、J_2 分别相应于参数 $\boldsymbol{\eta}$、$\boldsymbol{\gamma}$ 的梯度为

$$\begin{aligned}&{\rm grad}[J_1(\boldsymbol{\eta})]=\frac{\partial J_1(\boldsymbol{\eta})}{\partial\boldsymbol{\eta}}=\\ &-2\boldsymbol{\Phi}^{\rm T}(\hat{\boldsymbol{c}}_{k-1}\hat{\boldsymbol{d}}_{k-1},kT)[\boldsymbol{Y}(p,KT)-\\ &\boldsymbol{\Phi}(\hat{\boldsymbol{c}}_{k-1}\hat{\boldsymbol{d}}_{k-1},kT)\boldsymbol{\eta}],\\ &{\rm grad}[J_2(\boldsymbol{\gamma})]=\frac{\partial J_2(\boldsymbol{\gamma})}{\partial\boldsymbol{\gamma}}=\\ &-2\boldsymbol{\Psi}^{\rm T}(\hat{\boldsymbol{a}}_{k-1}\hat{\boldsymbol{b}}_{k-1},kT)[\boldsymbol{Y}(p,KT)-\\ &\boldsymbol{\Psi}(\hat{\boldsymbol{a}}_{k-1}\hat{\boldsymbol{b}}_{k-1},kT)\boldsymbol{\gamma}].\end{aligned}$$

根据随机梯度搜索原理, 极小化两个准则函数[14], 并考虑复杂度影响, 引入固定遗忘因子 λ_1 和可变遗忘因子 $\lambda_2(kT)$, 可得估计参数 $\hat{\boldsymbol{\eta}}$、$\hat{\boldsymbol{\gamma}}$ 的带可变遗忘因子的递阶多新息随机梯度递推算法 (VFF-HMISG) 为

$$\begin{aligned}&\hat{\boldsymbol{\gamma}}(kT)=\\ &\hat{\boldsymbol{\gamma}}(kT-T)+\frac{\boldsymbol{\Psi}^{\rm T}(\hat{\boldsymbol{a}}_{k-1}\hat{\boldsymbol{b}}_{k-1},kT)}{r_1(kT)}\times\\ &[\boldsymbol{Y}(p,kT)-\boldsymbol{\Psi}(\hat{\boldsymbol{a}}_{k-1}\hat{\boldsymbol{b}}_{k-1},kT)\hat{\boldsymbol{\gamma}}(kT-T)];\end{aligned} \tag{19}$$

$$\begin{aligned}&r_1(kT)=\lambda_1 r_1(kT-T)+||\psi(\hat{\boldsymbol{a}}_{k-1}\hat{\boldsymbol{b}}_{k-1},kT)||^2,\\ &r_1(0)=1,\ 0<\lambda_1\leqslant 1;\end{aligned} \tag{20}$$

$$\begin{aligned}&\hat{\boldsymbol{\eta}}(kT)=\\ &\hat{\boldsymbol{\eta}}(kT-T)+\frac{\boldsymbol{\Phi}^{\rm T}(\hat{\boldsymbol{c}}_{k-1}\hat{\boldsymbol{d}}_{k-1},kT)}{r_2(kT)}\times\\ &[\boldsymbol{Y}(p,kT)-\boldsymbol{\Phi}(\hat{\boldsymbol{c}}_{k-1}\hat{\boldsymbol{d}}_{k-1},kT)\hat{\boldsymbol{\eta}}(kT-T)];\end{aligned} \tag{21}$$

$$\begin{aligned}&r_2(kT)=\\ &\lambda_2(kT-T)r_2(kT-T)+||\varphi(\hat{\boldsymbol{c}}_{k-1}\hat{\boldsymbol{d}}_{k-1},kT)||^2,\\ &r_2(0)=1.\end{aligned} \tag{22}$$

其中 $1/r_1(kT)$、$1/r_2(kT)$ 为收敛因子. 遗忘因子用于调节收敛因子的大小, 进而影响对非平稳信号的自适应能力. 其值越小, 算法的跟踪能力越强, 但对于噪声

干扰也越敏感, 参数估计波动也越大; 反之, 对噪声干扰不敏感, 参数估计误差也较小, 但收敛速度较慢.

为了保证算法的收敛速度, 并提高算法的抗干扰能力, 减小参数估计误差, 本文 $\lambda_2(kT)$ 采用可变遗忘因子. 在参数辨识的初始阶段, 设定一个较小的 $\lambda_2(kT)$, 随着参数估计的稳定 $\lambda_2(kT)$ 逐渐趋向于 1, 由此提出一种新的遗忘因子修正函数

$$\lambda_2(kT)=\lambda_{\min}+(1-\lambda_{\min})^{2^{R(kT)}};$$

$$\text{s.t.}\ \begin{cases}R(kT)=\mathrm{NINT}(\rho|\xi(kT)|),\\ \xi(kT)=y(kT)-\hat{\boldsymbol{a}}_{k-1}^{\mathrm{T}}G(kT)\hat{\boldsymbol{d}}_{k-1}-\\ \qquad \hat{\boldsymbol{b}}_{k-1}^{\mathrm{T}}F(kT)\hat{\boldsymbol{c}}_{k-1}.\end{cases} \tag{23}$$

其中: $\lambda_2(kT)$ 为 kT 时刻的遗忘因子; $\lambda_{\min}$ 为遗忘因子最小值, $0\leqslant\lambda_{\min}\leqslant 1$; ρ 为敏感增益系数, 控制 $\lambda_2(kT)$ 趋向于 1 的速率; $\xi(kT)$ 为 kT 时刻输出量测值与估计值的误差; $\mathrm{NINT}(\cdot)$ 表示最接近 $(\cdot)$ 的整数.

由修正函数可知, 遗忘因子 $\lambda_2(kT)$ 由误差 $\xi(kT)$ 控制. 当 $\xi(kT)$ 较大时, $\lambda_2(kT)$ 较小, 算法能快速地跟踪信号变化; 反之, $\lambda_2(kT)$ 较大, 算法对噪声干扰不敏感, 收敛时参数估计误差也较小; 当 $\xi(kT)=0$ 时, $\lambda_2(kT)=1$ 可以保证参数估计误差最小.

在 VFF-HMISG 算法 (19) ∼ (23) 中:

1) 若 $\lambda_2(kT)$ 为常数, 则算法蜕变为递阶多新息遗忘随机梯度算法 (HMIFG);

2) 若 $p=1$, 则算法蜕变为带可变遗忘因子的递阶随机梯度算法 (VFF-HSG);

3) 若 $\lambda_{\min}=1$, 则算法蜕变为递阶多新息随机梯度算法 (HMISG);

4) 若 $p=1,\lambda_{\min}=1$, 则算法蜕变为递阶随机梯度算法 (HSG).

为叙述方便, 引入符号 $[\theta(kT)](p:q)$ 表示参数向量 $\theta(kT)$ 中第 $p\sim q$ 个向量数据, 则 VFF-HMISG 算法总结如下.

Step 1: 选择 p、$\lambda_{\min}$、ρ, 令 $k=1$, $\hat{\theta}(0)=\mathbf{1}_{n_a+mn_b+1}/p_0$, $\hat{\gamma}(0)=\mathbf{1}_{n_c+n_d}/p_0$, $p_0=10^6$, 这里 $\mathbf{1}_n$ 为元均为 1 的 n 维列向量.

Step 2: 采集 $u(kT+t_{j-1})$ 和 $y(kT)$, 由式 (1) 和 (2) 计算 $\bar{u}(kT+t_{j-1})$ 和 $\bar{y}(kT)$, 构造输入输出信息矩阵 $\boldsymbol{G}(kT)$ 和 $\boldsymbol{F}(kT)$. 再由式 (14)、(16)、(11)、(7)、(8) 分别构造堆积向量 $\boldsymbol{Y}(p,kT)$ 和广义信息向量 $\boldsymbol{\Psi}(\hat{\boldsymbol{a}}_{k-1}\hat{\boldsymbol{b}}_{k-1},kT)$.

Step 3: 由式 (20) 计算 $r_1(kT)$.

Step 4: 由式 (19) 刷新参数估计向量 $\hat{\boldsymbol{\gamma}}(kT)$, 并令

$$\hat{\boldsymbol{d}}_k:=[\hat{\gamma}(kT)](1:n_d),$$
$$\hat{\boldsymbol{c}}_k:=[\hat{\gamma}(kT)](n_d+1:n_d+n_c).$$

Step 5: 由式 (10) 构造广义信息向量 $\varphi(\hat{\boldsymbol{c}}_{k-1}\hat{\boldsymbol{d}}_{k-1},kT)$, 由式 (22) ∼ (23) 计算 $r_2(kT)$.

Step 6: 由式 (21) 刷新参数估计向量 $\hat{\boldsymbol{\eta}}(kT)$, 并令

$$\hat{\boldsymbol{a}}_k:=[\hat{\boldsymbol{\eta}}(kT)](1:n_a),$$
$$\hat{\boldsymbol{b}}_k:=[\hat{\boldsymbol{\eta}}(kT)](n_a+1:n_a+mn_b+1).$$

Step 7: $k=k+1$, 转 Step 2.

3 算法性能分析

3.1 计算量分析

为了评价算法的计算量, 用算法的乘法次数 (包括除法次数) 和加法次数 (包括减法次数) 来衡量. VFF-HMISG 算法每递推一步的乘法次数与加法次数分别为

$$3p(n_a+mn_b+1)+\min(n_a+mn_b+1,n_c+n_d)+p+1,$$
$$3p(n_a+mn_b+1)+\min(n_a+mn_b+1,n_c+n_d)+2,$$

则算法的计算量为

$$6p(n_a+mn_b+1)+2\min(n_a+mn_b+1,n_c+n_d)+p+3.$$

3.2 收敛性分析

由于在线性动态参数 $\hat{\boldsymbol{\eta}}(kT)$ 估计中引入了可变遗忘因子 $\lambda_2(kT)$, 其辨识过程更具一般性. 为此, 本文以 $\hat{\boldsymbol{\eta}}(kT)$ 为例, 分析算法的收敛性能.

引理 1 对于系统 (12) 及 VFF-HMISG 算法 (19) ∼ (23), 若信息向量 $\varphi(kT)$ 持续激励, 即存在常数 $0<\alpha\leqslant\beta<\infty$ 和正整数 $N\geqslant n$, 使得下述持续激励条件成立[15]:

$$\text{(A1)}\ \ \alpha\boldsymbol{I}\leqslant\frac{1}{N}\sum_{i=0}^{N-1}\varphi(kT+iT)\varphi^{\mathrm{T}}(kT+iT)\leqslant\beta\boldsymbol{I},\ \text{a.s.}\ \ k>0.$$

假设信息向量 $\varphi(kT)$ 有下界 $||\varphi(kT)||^2\geqslant\alpha>0$, 且 $r_2(0)$ 满足

$$r_2(0)\leqslant\frac{nN\beta}{1-\lambda_{\min}},\ 0\leqslant\lambda_{\min}<1, \tag{24}$$

则对于所有的 $k>0$, 式 (22) 中的 $r_2(kT)$ 满足不等式

$$r_2(kT)\leqslant\frac{nN\beta}{1-\lambda_{\min}},\ \text{a.s.}\ 0\leqslant\lambda_{\min}<1. \tag{25}$$

定理 1 对于系统 (12) 以及 VFF-HMISG 算法 (19) ∼ (23), 若条件 (A1) 成立, 则 $r_2(0)$ 由式 (24) 确定, 且观测噪声 $\{v(kT)\}$ 为满足条件

$$\text{(A2)}\qquad E[v(kT)]=0,$$
$$\text{(A3)}\qquad E[v^2(kT)]\leqslant\sigma_v^2<\infty$$

的零均值随机序列. 令新息长度 $p=N$, $E[||\tilde{\eta}(0)||^2]=E[||\hat{\eta}(0)-\eta(0)||^2]=\delta_0<\infty$. 则 VFF-HMISG 算法的参数估计误差是均方有界的, 即

$$E[||\hat{\eta}(kT)-\eta(kT)||^2]\leqslant$$
$$[\sqrt{1-\mu}]^k\delta_0+\frac{1}{(1-\sqrt{1-\mu})^2}\times$$

$$\frac{N^2\beta(1-\lambda_{\min})^2\sigma_v^2}{\alpha^2}=:f_u(\lambda_{\min},k).$$

其中: $0<\mu:=\dfrac{\alpha(1-\lambda_{\min})}{n\beta}<1,\ 0<\lambda_{\min}<1,\ 0<\mu<1$, δ_0 为有限值, 当 k 趋于无穷大时, $[\sqrt{1-\mu}]^k\delta_0$ 趋于 0. 从而得到参数估计误差的近似上界为

$$E[||\hat{\eta}(kT)-\eta(kT)||^2]=f_u(\lambda_{\min},k)\approx$$
$$\frac{1}{(1-\sqrt{1-\mu})^2}\times\frac{N^2\beta(1-\lambda_{\min})^2\sigma_v^2}{\alpha^2}=$$
$$f(a_0,\lambda_{\min})=:f(\lambda_{\min}).$$

文献 [16] 给出了引理 1 和定理 1 的证明过程, 本文不再赘述.

4　仿真验证

考虑某一非均匀采样数据 H-W 非线性系统

$$y(kT)=$$
$$\alpha(z)\sum_{l=1}^{n_d}d_lg_l(y(kT))+$$
$$\sum_{j=1}^{m}\beta_j(z)\sum_{s=1}^{n_c}c_sf_s(u(kT+t_{j-1}))+v(kT),\qquad(26)$$
$$\alpha(z)=\alpha_1z^{-1}+\alpha_2z^{-2}=-1.529z^{-1}+10.5z^{-2},$$
$$\beta_1(z)=\beta_{10}+\beta_{11}z^{-1}+\beta_{12}z^{-2}=$$
$$0.123\,4+0.068\,99z^{-1}+0.015\,38z^{-2},$$
$$\beta_2(z)=\beta_{21}z^{-1}+\beta_{22}z^{-2}=$$
$$0.042\,1z^{-1}+0.085\,06z^{-2},$$
$$\bar{u}(kT+t_i)=\sum_{s=1}^{n_c}c_sf_s(u(kT+t_i))=$$
$$0.6\sin(u(kT+t_i))+0.1\cos(u^2(kT+t_i)),$$
$$\bar{y}(kT)=\sum_{l=1}^{n_d}d_lg_l(y(kT))=$$
$$0.2\sin(y(kT))+0.1\cos(y^2(kT)),$$
$$\Theta=[\alpha_1,\alpha_2,\beta_{10},\beta_{11},\beta_{12},\beta_{21},\beta_{22},c_1,c_2]^{\rm T}.$$

系统输入非均匀刷新次数 $m=2$, 输出采样周期 $T=2.5\,{\rm s}$, 输入采样间隔 $\tau_1=1\,{\rm s}$, $\tau_2=1.5\,{\rm s}$, 则 $t_0=0\,{\rm s}$, $t_1=\tau_1=1\,{\rm s}$, $t_2=\tau_1+\tau_2=T=2.5\,{\rm s}$, 输入 $\{u(kT),u(kT+t_1)\}$ 以零均值单位方差持续信号激励系统, $\{v(kT)\}$ 为零均值方差为 σ^2 的白噪声, 数据长度设为 4 000. 采用 VFF-HMISG 算法辨识模型 (26) 的参数. 为了评估算法的有效性, 定义性能指标如下:

$$\delta=\left\|\frac{\hat{\theta}(kT)-\theta}{\theta}\right\|\times100\%.\qquad(27)$$

当 $\sigma^2=0.1$ 时, 选取 $p=2$, $\lambda_{\min}=0.9$, $\rho=5$, $\lambda_1=1$, 参数估计及不同采样周期长度下的估计误差 δ 如表 1 所示. 由表 1 可知, 随着采样周期的增加, 参数估计误差减小, 到 1 000 周期时, 趋于稳定.

表 1　VFF-HMISG 估计及误差

k	100	200	500	1 000	2 000	3 000	4 000	真值
a_1	−1.473 77	−1.538 48	−1.573 61	−1.532 66	−1.528 49	−1.514 77	−1.514 34	−1.529 00
a_2	10.249 50	10.386 04	10.388 51	10.465 99	10.481 63	10.492 37	10.496 64	10.500 00
b_{10}	−0.004 26	−0.004 55	−0.005 02	−0.005 41	−0.005 66	−0.005 64	−0.005 64	0.123 40
b_{11}	−0.012 09	−0.012 54	−0.012 97	−0.013 18	−0.013 53	−0.013 64	−0.013 65	0.068 99
b_{12}	−0.009 35	−0.009 56	−0.010 01	−0.010 32	−0.010 63	−0.010 79	−0.010 88	0.015 38
b_{21}	0.006 58	0.006 37	0.005 88	0.005 43	0.005 20	0.005 05	0.005 02	0.042 10
b_{22}	0.0082 8	0.008 17	0.008 23	0.007 82	0.007 50	0.007 44	0.007 42	0.085 06
d_1	−0.199 53	−0.200 66	−0.199 53	−0.200 09	−0.200 36	−0.200 49	−0.200 47	0.200 00
d_2	−0.101 11	−0.101 29	−0.099 51	−0.099 80	−0.099 45	−0.099 39	−0.099 52	0.100 00
c_1	0.000 09	0.000 14	0.000 24	0.000 32	0.000 39	0.000 44	0.000 48	0.600 00
c_2	−0.001 41	−0.001 38	−0.001 32	−0.001 26	−0.001 19	−0.001 17	−0.001 15	0.100 00
$\delta/\%$	7.679 76	7.376 52	7.375 48	7.299 84	7.295 89	7.295 76	7.295 57	0.000 00

参数 p、$\lambda_{\min}$、ρ 对算法的收敛速度与精度的影响较大, 其性能指标变化曲线如图 2 所示. 可以看出: 适当增大新息长度, 算法的辨识性能有所提高; 引入可变遗忘因子后, 算法的收敛速度明显加快 (如图 2 中虚线所示), 且 800 周期后参数估计较为平稳; 适当改变增益 ρ 的取值, 参数估计的收敛速度更为敏感 (如图 2 中点线所示).

为了验证算法对噪声的抗干扰能力, 选取 $p=2$, $\lambda_1=1$, σ^2 分别为 0.1、0.3 时 VFF-HMISG ($\lambda_{\min}=0.95$, $\rho=5$)、HMIFG ($\lambda_2=0.95$) 算法的估计误差随采样周期长度的变化曲线如图 3 所示.

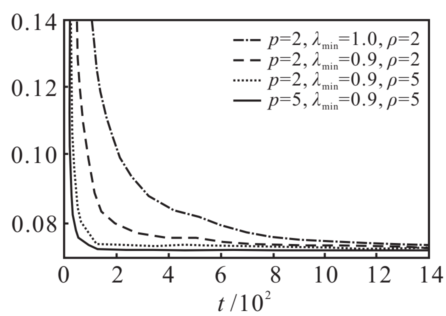
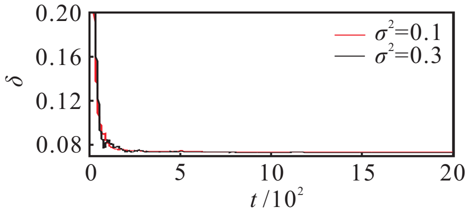
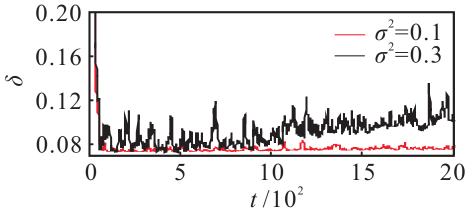


图 2　VFF-HMISG 参数估计误差 δ 变化曲线



(a) VFF-HMISG算法的估计性能



(b) HMISG算法的估计性能

图 3 不同方差下两种算法的参数估计误差 δ 变化曲线

可以看出, 在辨识过程中, 噪声扰动增大时, 算法 HMIFG 出现了明显波动, 而 VFF-HMISG 算法在保证辨识精度的同时, 具有更强的抗干扰能力.

5 结 论

本文针对非均匀采样数据 H-W 非线性系统, 提出了带可变遗忘因子的递阶多新息随机梯度辨识算法. 相比文献 [10] 中先辨识转换模型参数, 后分离出系统参数的方法, 本文算法直接辨识系统参数, 降低了计算量, 引入可变遗忘因子后, 算法收敛速度提高, 对噪声的抗干扰能力有所增强. 该算法经简单的变换即可直接用于 Hammerstein、Wiener 系统的参数辨识, 并可结合其他参数辨识方法, 用于线性及非线性系统的参数估计、状态识别及故障检测, 但 p、$\lambda_{\min}$、ρ 选择等问题还有待进一步研究.

参考文献(References)

[1] 丁洁, 谢莉, 丁锋. 非均匀采样系统多新息随机梯度辨识性能分析[J]. 控制与决策, 2011, 26(9): 1338-1342.
(Ding J, Xie L, Ding F. Performance analysis of multi-innovation stochastic gradient identification for non-uniformly sampled systems[J]. Control and Decision, 2011, 26(9): 1338-1342.)

[2] Liu Y, Ding F, Shi Y. Least squares estimation for a class of non-uniformly sampled systems based on the hierarchical identification principle[J]. Circuits, Systems, and Signal Processing, 2012, 31(6): 1985-2000.

[3] 蒋红霞, 王金海, 丁锋. 一类非均匀采样系统最小二乘迭代辨识[J]. 系统工程与电子技术, 2008, 30(8): 1535-1539.
(Jiang H X, Wang J H, Ding F. Least-square-iterative identification of a class of non-uniform sampled-data systems[J]. Systems Engineering and Electronics, 2008, 30(8): 1535-1539.)

[4] Xie L, Yang H. Gradient-based iterative identification for nonuniform sampling output error systems[J]. J of Vibration and Control, 2011, 17(3): 471-478.

[5] 谢莉, 丁锋. 非均匀采样数据系统的一种辨识方法[J]. 控制工程, 2008, 15(4): 402-404.
(Xie L, Ding F. Identification method of non-uniformly sampled-data systems[J]. Control Engineering of China, 2008, 15(4): 402-404.)

[6] 谢莉, 刘艳君. 非均匀采样数据系统 AM-MI-GESG 算法[J]. 仪器仪表学报, 2009, 30(6): 25-29.
(Xie L, Liu Y J. AM-MI-GESG algorithms for non-uniformly sampled-data systems[J]. Chinese J of Scientific Instrument, 2009, 30(6): 25-29.)

[7] Xie L, Yang H Z, Ding F. Recursive least squares parameter estimation for non-uniformly sampled systems based on the data filtering[J]. Mathematical and Computer Modeling, 2011, 54(1/2): 315-324.

[8] 陈聪. 基于 Hammerstein 模型的双速率系统辨识[D]. 杭州: 浙江大学控制科学与工程学系, 2010.
(Chen C. Nonlinear identification of dual-rate systems based on Hammerstein model[D]. Hangzhou: Department of Control Science and Engineering, Zhejiang University, 2010.)

[9] Jing Chen, Lixing Lü, Ruifeng Ding. Multi-innovation stochastic gradient algorithms for dual-rate sampled systems with preload nonlinearity[J]. Applied Mathematics Letters, 2013, 26(1): 124-129.

[10] Chen J, Lü L X, Ding R F. Parameter estimation for dual-rate sampled data systems with preload nonlinearities[J]. Advances in Intelligent and Soft Computing, 2011, 125: 43-50.

[11] Wang D Q, Ding F. Least squares based and gradient based iterative identification for Wiener nonlinear systems[J]. Signal Processing, 2011, 91(5): 1182-1189.

[12] Ding F. Hierarchical multi-innovation stochastic gradient algorithm for Hammerstein nonlinear system modeling[J]. Applied Mathematical Modelling, 2013, 37(4): 1694-1704.

[13] Ding F, Qiu L, Chen T. Reconstruction of continuous-time systems from their non-uniformly sampled discrete-time systems[J]. Automatica, 2009, 45(2): 324-332.

[14] Goodwin G C, Sin K S. Adaptive filtering prediction and control[M]. Englewood Cliffs: Prentice-hall, 1984: 68-79.

[15] Ljung L. Consistency of the least-squares identification method[J]. IEEE Trans on Automatic Control, 1976, 21(5): 779-781.

[16] Ding F, Chen T W. Performance analysis of multi-innovation gradient type identification methods[J]. Automatica, 2007, 43(1): 1-14.

(责任编辑：李君玲)