

双率系统辅助模型框架下的随机牛顿递推辨识

曹鹏飞, 罗雄麟

(中国石油大学 自动化研究所, 北京 102249)

摘要: 针对双率系统, 采用基于辅助模型的改进随机牛顿递推算法辨识输出误差模型. 若当前参数估计对应的估计系统不稳定, 则出现中间不可测时刻输出估计发散, 辨识过程停止. 增加非线性模型与常规辅助模型一起为下步递推提供信息估计, 确保递推进行. 为避免出现输入不充分或者广泛时 Hessian 阵奇异或者接近奇异的情况, 在 Hessian 阵的递推中增加对称正定矩阵. 最后给出了所提出辨识算法的一致收敛性证明.

关键词: 双率系统; 辨识; 双辅助模型; 收敛分析; 改进随机牛顿算法

中图分类号: TP273

文献标志码: A

Auxiliary-model-based stochastic Newton recursive identification for dual-rate system

CAO Peng-fei, LUO Xiong-lin

(Research Institute of Automation, China University of Petroleum, Beijing 102249, China. Correspondent: LUO Xiong-lin, E-mail: luoxl@cup.edu.cn)

Abstract: An auxiliary-model-based improved stochastic Newton recursive algorithm is utilized to identify the output-error model for the dual-rate system. The output estimations at the time between two slow-sampled periods will diverge when the unstable estimated system arises and the identification process will cease. Therefore, a nonlinear model is proposed as the sub-auxiliary model. Together with the general auxiliary model, the proposed model provides estimated information for the next recursion. When the input is not general enough, the Hessian matrix will be singular or nearly singular. A positive definite symmetric matrix is added to the recursion for Hessian matrix to ensure it positive definite. Finally, the uniform convergence of the proposed algorithm is proved.

Keywords: dual-rate system; identification; double auxiliary models; convergence analysis; improved stochastic Newton algorithm

0 引言

由于受到客观物理因素的限制, 连续系统的输入输出量有不同的采样频率, 从而形成了多采样频率系统(多率系统)^[1]. 双率系统是一种特殊的多率系统, 其广泛存在, 并具有较强的研究意义. 双率系统控制不仅在石油化工过程控制方面得到了成功应用, 在理论方面也取得了一系列研究成果, 例如双率系统自适应控制^[2]、双率系统预测控制^[3]、鲁棒控制^[4]等. 而能否得到双率模型则是控制研究首要解决的问题^[5-6], 因此双率系统辨识成为双率系统研究的最主要的热点和难点之一. 一般是针对某种模型结构采用某些辨识算法, 例如基于状态空间模型利用提升

技术辨识^[7]、基于小波模型利用小波多分辨率分析辨识^[8]、基于差分方程利用多项式转换技术辨识^[9-10]等等.

本文主要针对快采样输入和慢采样输出的双率系统建立输出误差模型, 并基于此模型研究辨识算法. 多项式转换技术^[9-10]是辨识双率系统模型的主要方法, 单率模型经过多项式技术转化为仅依赖于直接测量数据辨识的双率模型, 并利用最小二乘算法辨识双率模型参数和估计不可测时刻输出. 然而, 伴随更多参数的出现, 计算量成倍增加, 因此多项式技术也存在一定的局限性. Ding 等^[11-12]在多项式变换后, 首先根据递阶原理将双率模型分成多个维数较小的子系

收稿日期: 2013-09-05; 修回日期: 2014-03-05.

基金项目: 国家973计划项目(2012CB720500); 国家自然科学基金项目(61104218); 中国石油大学(北京)科研基金项目(YJRC-2013-12).

作者简介: 曹鹏飞(1988—), 男, 博士生, 从事软测量、系统辨识的研究; 罗雄麟(1963—), 男, 教授, 博士生导师, 从事控制理论、过程控制工程、化工系统工程、机器学习等研究.

统,并分别辨识各个子系统和调整子系统间的关联项.虽然这种递阶分解的方式能够降低各个子系统的辨识维数,但仍无法避免大周期比(输出与输入的采样周期比值)下辨识的计算量问题. Raghavan 等^[13]提出了期望最大化迭代算法,根据可测数据辨识得到采样间值,避免了利用多项式转换技术出现大量参数的问题.然而,这种插值方法只能提供一定可靠性的中间时刻插值,并不能完全解决大周期比情况下的辨识问题.丁锋等^[14-16]提出了基于辅助模型的辨识算法,该方法基于可测信息建立辅助模型,利用辅助模型输出代替系统模型不可测变量.通过辅助模型参变量的调整,使得辅助模型的输出逼近不可测变量,最终得到系统模型参数一致收敛.

基于辅助模型的辨识方法避免了双率模型高维参数辨识的问题,能够给出较好的参数一致估计,因此本文利用已有方法建立辅助模型,并基于辅助模型利用递推辨识算法估计系统模型参数.为避免中间不可测时刻输出估计发散影响参数收敛,提出增加非线性模型作为次辅助模型,与原有辅助模型(主辅助模型)一起为下步递推提供所需信息估计.另外,递推算法采用改进随机牛顿算法,在原 Hessian 矩阵递推中增加了对称正定矩阵,保证 Hessian 阵在输入不充分或者广泛的情况下正定.最后给出了基于双辅助模型的改进随机牛顿递推辨识算法一致收敛性证明.

1 双辅助模型与改进随机牛顿算法的必要性分析

基于常规辅助模型方法辨识双率系统模型参数,如图 1 虚线框 1 所示.当周期比较大时,如果在某一步递推获得的参数估计对应的估计系统是不稳定的,则容易出现基于辅助模型的中间不可测时刻输出估计的发散,影响了递推的进行和参数的收敛.这个问题在最小二乘和随机梯度等经典递推算法应用在双率辨识中同样会出现.为了解决该问题,本文提出增加辅助模型的方法,即在原有辅助模型(主辅助模型)基础上增加一个非线性模型作为次辅助模型,如图 1 虚线框 2 所示.两个辅助模型之间有一个切换机制,当主辅助模型中间不可测时刻的输出估计出现发散,则切换到次辅助模型,利用次辅助模型给出下步递推所需信息估计.这样的方式保证了递推辨识的进行,最终获得参数估计一致收敛.

本文采用改进随机牛顿递推算法辨识模型.随机牛顿算法更加接近递推辨识算法的一般形式,并且以它为基础可以揭示很多辨识算法的内在联系,如最小二乘递推算法、梯度递推算法等^[17].因此,随机牛顿算法是一种更加有研究价值的算法形式.然而, Ljung^[18]指出,当输入信号并不充分或者广泛时容易

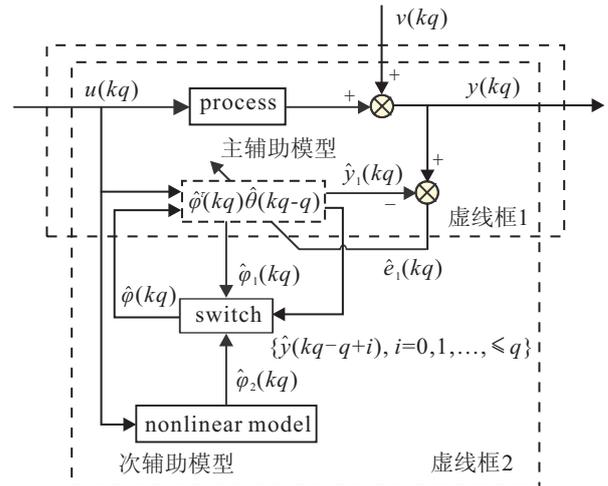


图 1 递推辨识算法流程

出现 Hessian 阵奇异或者接近奇异的情况,会影响到参数的递推过程.因此这里提出改进,在 Hessian 阵的递推中增加一个对称正定矩阵,该矩阵在满足一定条件下保证递推的有效进行,并且不影响参数估计一致收敛.

2 基于双辅助模型的改进随机牛顿递推算法

2.1 改进随机牛顿递推算法

一般线性时不变系统输出误差模型可用下式表示:

$$\begin{aligned} A(z)y(k) &= B(z)u(k) + v(k); \\ A(z) &= 1 + a_1z^{-1} + \cdots + a_{n_a}z^{-n_a}, \\ B(z) &= b_1z^{-1} + \cdots + b_{n_b}z^{-n_b}. \end{aligned}$$

整理得到

$$\begin{aligned} y(k) &= \varphi^T(k)\theta + v(k); \\ \varphi^T(k) &= [-y(k-1), -y(k-2), \cdots, -y(k-n_a), \\ &\quad u(k-1), \cdots, u(k-n_b)], \\ \theta &= [a_1, a_2, \cdots, a_{n_a}, b_1, b_2, \cdots, b_{n_b}]^T. \end{aligned} \quad (1)$$

其中: $y(k)$ 为输出变量, $u(k)$ 为输入变量, $v(k)$ 为随机噪声, $\varphi(k)$ 被称为信息.这里采用改进随机牛顿递推算法辨识模型.针对系统(1),具体随机牛顿参数递推算法如下所示:

$$\begin{cases} \hat{\theta}(k) = \hat{\theta}(k-1) + \rho(k)R^{-1}(k)\varphi(k)[y(k) - \\ \quad \varphi^T(k)\hat{\theta}(k-1)], \\ R(k) = R(k-1) + \rho(k)[\varphi(k)\varphi^T(k) - \\ \quad R(k-1)]. \end{cases}$$

其中: $\rho(k)$ 为收敛因子, $R(k)$ 为 Hessian 阵.在 $R(k)$ 的递推公式中增加一个对称正定矩阵 $A(k)$,其在满足一定条件下不影响参数估计的收敛结果,具体条件在后面收敛性证明中给出.改进随机牛顿递推算法如下所

示:

$$\begin{cases} \hat{\theta}(k) = \hat{\theta}(k-1) + \rho(k)R^{-1}(k)\varphi(k)[y(k) - \\ \varphi^T(k)\hat{\theta}(k-1)], \\ R(k) = R(k-1) + \rho(k)[\varphi(k)\varphi^T(k) + \\ A(k) - R(k-1)]. \end{cases} \quad (2)$$

2.2 双辅助模型

通常可以利用式(2)的方法获得系统(1)的模型参数,然而对于双率系统而言,输出变量的采样周期大于输入变量采样周期(本文规定周期比为 q ,即 $T_y = qT_u$),无法获取中间不可测时刻数据,因此式(2)递推算法不能直接应用于双率系统参数估计.本文基于双辅助模型方法辨识双率系统模型.

针对系统(1),利用上一步估计的参数获得信息估计 $\hat{\varphi}(kq)$,并代替信息 $\varphi(kq)$ 进行下一步递推.基于辅助模型的改进随机牛顿递推算法如下所示:

$$\begin{cases} \hat{\theta}(kq) = \hat{\theta}(kq-q) + \rho(kq)R^{-1}(kq)\hat{\varphi}(kq)[y(kq) - \\ \hat{\varphi}^T(kq)\hat{\theta}(kq-q)], \\ R(kq) = R(kq-q) + \rho(kq)[\hat{\varphi}(kq)\hat{\varphi}^T(kq) + \\ A(kq) - R(kq-q)]. \end{cases} \quad (3)$$

其中

$$\hat{\varphi}^T(kq) = [-\hat{y}(kq-1), -\hat{y}(kq-2), \dots, -\hat{y}(kq-n_a), \\ u(kq-1), \dots, u(kq-n_b)],$$

$$\hat{y}(kq-i) = \hat{\varphi}^T(kq-i)\hat{\theta}(kq-q), \quad i = 1, 2, \dots, q-1.$$

辅助模型为 $\hat{y}(kq) = \hat{\varphi}^T(kq)\hat{\theta}(kq-q)$,本文称为主辅助模型. $\hat{y}(kq-i)$ 为该段慢采样周期内中间不可测输出估计.

通常在实际应用中 $A(kq)$ 的选取较麻烦,为此利用 $\lambda\mathbf{I}$ 代替 $A(kq)$,而在收敛性证明中仍然采用 $A(kq)$.因此有

$$\begin{cases} \hat{\theta}(kq) = \hat{\theta}(kq-q) + \rho(kq)R^{-1}(kq)\hat{\varphi}(kq)[y(kq) - \\ \hat{\varphi}^T(kq)\hat{\theta}(kq-q)], \\ R(kq) = R(kq-q) + \rho(kq)[\hat{\varphi}(kq)\hat{\varphi}^T(kq) + \\ \lambda\mathbf{I} - R(kq-q)]. \end{cases} \quad (4)$$

为了避免发散现象影响递推算法进行,本文提出利用非线性模型作为常规辅助模型补充,为递推算法提供信息估计.Khargonekar^[19]认为由于双率的存在,使得慢采样周期的采样点表现出了周期时变的特性,而这种特性完全可以利用非线性模型来体现,因此可以利用慢采样数据建立非线性模型,并给出一定可靠性的输出估计.另外,在实际工业过程中软测量技术是重要的检测手段,其核心是软测量模型,而软测量

模型的输入输出则具有不同的采样周期,大多建模方法依赖于这些慢采样数据建立非线性软测量模型,并且得到了较好的应用效果^[20-21].因此建立这样一个非线性模型作为次辅助模型是可信的,并且可以借鉴软测量建模的方法建立本文的次辅助模型,模型的具体形式不为研究重点,在此不过多描述.

这里有4点说明:

1) 设定阈值判断是否发散,当中间时刻的输出估计不超过这个阈值,则使用主辅助模型为下步递推提供信息估计,否则选择次辅助模型提供信息估计,而这个阈值的确定根据具体输出情况而定,并且这个阈值通常会影响到辅助模型选择情况以及参数估计的收敛过程.

2) 不单用非线性模型作为辅助模型的原因在于,一般系统是处于动态过程,而非线性模型的建立依赖于慢采样数据,因此无法表征系统动态特性,对输出的估计只具有一定的可靠性.非线性模型仅在递推过程中提供有界信息估计,保证递推进行;而主辅助模型则会在参数逐步逼近真实值时,接近中间不可测时刻输出真值,进而促进参数估计收敛.在实际应用中,一般只在递推辨识前期会出现发散的现象,在后面的辨识过程中基本只依赖主辅助模型.

3) 当出现发散现象并采用上一步稳定的递推结果作为当前的参数估计,同时预测该段慢采样周期内中间时刻输出,都不能确保以后不会出现参数估计不稳定的结果,而且若采用这样的方式,很有可能会出现参数估计收敛在某个参数值而无法更新的情况,这个现象在仿真实验中曾出现过.

4) 利用非线性模型估计中间时刻的输出类似于插值,然而,不同的是非线性模型首先只是在必要的时候才会提供信息估计;其次,慢采样数据样本间的关系可以利用非线性模型表征,并且这种方法已经在实际中成功应用,因此利用非线性模型提供信息估计是可行的.信息估计值对辨识过程会有很大影响,但对递推辨识结果的影响相对较小,这点可从下面的证明中看出.利用非线性模型提供信息估计具有较高的可靠性.

2.3 递推辨识步骤

具体递推辨识流程如图1所示.假设有 M 个慢采样数据,且样本数据充分.具体递推步骤如下:

1) 基于 M 个慢采样数据建立非线性模型,如神经网络、支持向量机、多项式模型等,并根据具体情况设定阈值.

2) 令 $\hat{\theta}(0) = \varepsilon$, $R(0) = a^2\mathbf{I}$,其中 ε 为充分小的实向量, a 为充分大的实数.

3) 利用 $\hat{\theta}(kq - q)$ 估计下一慢采样周期之前的中间时刻输出, 若出现发散现象, 则选择利用非线性模型得到 $\hat{\varphi}(kq)$, 并利用式 (4) 获得下一步参数递推结果.

4) 重复 3), 直到参数收敛.

3 改进随机牛顿双率辨识递推算法收敛性分析

首先给出一个定义: 对于任意函数 $f(x) \geq 0$, $g(x) \geq 0$, 定义 $f(x) = O(g(x))$, 表征存在正数 γ 满足 $f(x) \leq \gamma g(x)$.

本文收敛性分析所用假设条件如下:

A1: 收敛因子满足

$$\sum_{i=1}^{\infty} \rho(iq) = \infty, \quad \sum_{i=1}^{\infty} \rho^2(iq) < \infty.$$

A2: 噪声满足

$$\begin{aligned} E[v(kq) | F_{kq-q}] &= 0, \text{ a.s.}, \\ E[v^2(kq) | F_{kq-q}] &\leq \sigma_v^2 < \infty, \text{ a.s.}, \end{aligned}$$

其中 F_{kq-q} 是由直到 $kq - q$ 时刻的观测生成的 σ -代数.

A3: $\{A(kq)\}$ 都满足对称正定, 且

$$\frac{\lambda_{\max}(A(kq))}{\lambda_{\min}(R(kq - q))} = 1,$$

或

$$\frac{\lambda_{\max}(A(kq))}{\delta} = 1,$$

其中 δ 为满足 $\delta \leq \lambda_{\min}(R(kq - q))$ 的充分小的正数.

A4: $A(z)$ 满足 $\frac{2 - A(z)}{A(z)}$ 严格正实.

A5: 信息估计满足

$$0 < \hat{\varphi}^T(kq)\hat{\varphi}(kq) \leq m < \infty.$$

这个条件很容易满足, 利用非线性模型估计不可测时刻输出值, 可保证信息向量有界, 则上述条件得到保证.

A6: 信息估计在一定输入信号下满足

$$\alpha \mathbf{I} \leq \frac{1}{L} \sum_{j=0}^{L-1} \hat{\varphi}(iq + jq)\hat{\varphi}^T(iq + jq) \leq \beta \mathbf{I},$$

其中 $\beta \geq \alpha > 0$.

引理 1 (鞅收敛定理^[23]) 若 $\{X(t)\}$ 是适应于 σ 代数 $\{F_t\}$ 增序列的非负随机变量序列, 并且

$$E[X(t+1) | F_t] \leq X(t) - a(t) + b(t), \text{ a.s.},$$

$E[X(0)] < \infty$, 且 $\sum_{i=1}^{\infty} b(i) < \infty$, 则 $X(t)$ 几乎收敛于有穷随机变量, 并且 $\sum_{i=1}^{\infty} a(i) < \infty$. 其中: $a(t) \geq 0$, $b(t) \geq 0$.

引理 2 在满足 A1、A3、A5 条件下, 从式 (3) 易

知 $R(kq)$ 总是有界对称正定矩阵, 并且存在正整数 N , 使得当 $k \geq N$ 时 $R(kq) \rightarrow R$, 其中 R 为某一有界对称正定矩阵.

定理 1 在满足 A1 ~ A5 的条件下, 本文提出的递推方法可保证双率系统模型参数估计一致有界.

证明 定义

$$e(kq) = y(kq) - \hat{\varphi}^T(kq)\hat{\theta}(kq - q),$$

$$\eta(kq) = y(kq) - \hat{\varphi}^T(kq)\hat{\theta}(kq),$$

易得

$$\eta(kq) = (1 - \rho(kq)\hat{\varphi}^T(kq)R^{-1}(kq)\hat{\varphi}(kq))e(kq).$$

令

$$\omega(kq) = 1 - \rho(kq)\hat{\varphi}^T(kq)R^{-1}(kq)\hat{\varphi}(kq),$$

有

$$\eta(kq) = \omega(kq)e(kq). \quad (5)$$

另外, 定义 $V(kq) = \tilde{\theta}^T(kq)R(kq)\tilde{\theta}(kq)$, 则有

$$\begin{aligned} V(kq) &= \\ &\{\tilde{\theta}(kq - q) + \rho(kq)R^{-1}(kq)\hat{\varphi}(kq)e(kq)\}^T \times \\ &R(kq)\{\tilde{\theta}(kq - q) + \rho(kq)R^{-1}(kq)\hat{\varphi}(kq)e(kq)\} = \\ &\tilde{\theta}^T(kq - q)R(kq)\tilde{\theta}(kq - q) + \\ &2\rho(kq)\tilde{\theta}^T(kq - q)\hat{\varphi}(kq)e(kq) + \\ &\rho^2(kq)\hat{\varphi}^2(kq)R^{-1}(kq)\hat{\varphi}(kq)e^2(kq) = \\ &(1 - \rho(kq))\tilde{\theta}^T(kq - q)R(kq - q)\tilde{\theta}(kq - q) + \\ &\rho(kq)\tilde{\theta}^T(kq - q)A(kq)\tilde{\theta}(kq - q) + \\ &\rho(kq)(\tilde{\theta}^T(kq - q)\hat{\varphi}(kq))^2 + \\ &2\rho(kq)\tilde{\theta}^T(kq - q)\hat{\varphi}(kq)e(kq) + \\ &\rho^2(kq)\hat{\varphi}^2(kq)R^{-1}(kq)\hat{\varphi}(kq)e^2(kq) \leq \\ &\left(1 - \rho(kq) + \rho(kq)\frac{\lambda_{\max}(A(kq))}{\lambda_{\min}(R(kq - q))}\right) \times \\ &\tilde{\theta}^T(kq - q)R(kq - q)\tilde{\theta}(kq - q) + \\ &\rho(kq)(\tilde{\theta}^T(kq - q)\hat{\varphi}(kq))^2 + \\ &2\rho(kq)\tilde{\theta}^T(kq - q)\hat{\varphi}(kq)e(kq) + \\ &\rho^2(kq)\hat{\varphi}^2(kq)R^{-1}(kq)\hat{\varphi}(kq)e^2(kq). \end{aligned}$$

由 A3 可知

$$\begin{aligned} V(kq) &\leq \\ &V(kq - q) + \rho(kq)(\tilde{\theta}^T(kq - q)\hat{\varphi}(kq))^2 + \\ &2\rho(kq)\tilde{\theta}^T(kq - q)\hat{\varphi}(kq)e(kq) + \\ &\rho^2(kq)\hat{\varphi}^2(kq)R^{-1}(kq)\hat{\varphi}(kq)e^2(kq) = \\ &V(kq - q) + \rho(kq)(\tilde{\theta}^T(kq - q)\hat{\varphi}(kq))^2 + \\ &\rho^2(kq)\hat{\varphi}^2(kq)R^{-1}(kq)\hat{\varphi}(kq)e^2(kq) + 2\rho(kq)(\tilde{\theta}(kq) - \\ &\rho(kq)R^{-1}(kq)\hat{\varphi}(kq)e(kq))^T \hat{\varphi}(kq)e(kq) = \end{aligned}$$

$$\begin{aligned}
& V(kq - q) + \rho(kq)(\tilde{\theta}^T(kq - q)\hat{\varphi}(kq))^2 - \\
& \rho^2(kq)\hat{\varphi}^2(kq)R^{-1}(kq)\hat{\varphi}(kq)e^2(kq) + \\
& 2\rho(kq)\tilde{\theta}^T(kq)\hat{\varphi}(kq)e(kq) = \\
& V(kq - q) + \rho(kq)(\tilde{\theta}^T(kq - q)\hat{\varphi}(kq))^2 - \\
& \rho^2(kq)\hat{\varphi}^2(kq)R^{-1}(kq)\hat{\varphi}(kq)e^2(kq) + \\
& 2\rho(kq)\tilde{\theta}^T(kq)\hat{\varphi}(kq)\omega^{-1}(kq)(\eta(kq) - v(kq)) + \\
& 2\rho(kq)\tilde{\theta}^T(kq)\hat{\varphi}(kq)\omega^{-1}(kq)v(kq). \quad (6)
\end{aligned}$$

其中

$$\begin{aligned}
& 2\rho(kq)\tilde{\theta}^T(kq)\hat{\varphi}(kq)\omega^{-1}(kq)v(kq) = \\
& 2\rho(kq)(\tilde{\theta}(kq - q) + \\
& \rho(kq)R^{-1}(kq)\hat{\varphi}(kq)e(kq))^T\hat{\varphi}(kq)\omega^{-1}(kq)v(kq) = \\
& 2\rho(kq)\tilde{\theta}^T(kq - q)\hat{\varphi}(kq)\omega^{-1}(kq)v(kq) + \\
& 2\rho^2(kq)\hat{\varphi}^T(kq)R^{-1}(kq)\hat{\varphi}(kq)\omega^{-1}(kq)v^2(kq) + \\
& 2\rho^2(kq)\hat{\varphi}^T(kq)R^{-1}(kq)\hat{\varphi}(kq)\omega^{-1}(kq)(e(kq) - \\
& v(kq))v(kq). \quad (7)
\end{aligned}$$

另外

$$\begin{aligned}
& \rho(kq)(\tilde{\theta}^T(kq - q)\hat{\varphi}(kq))^2 = \\
& \rho(kq)(\tilde{\theta}(kq) - \rho(kq)R^{-1}(kq)\hat{\varphi}(kq)e(kq))^T\hat{\varphi}(kq) \times \\
& \hat{\varphi}^T(kq)(\tilde{\theta}(kq) - \rho(kq)R^{-1}(kq)\hat{\varphi}(kq)e(kq)) = \\
& \rho(kq)(\tilde{\theta}^T(kq)\hat{\varphi}(kq))^2 - \\
& 2\rho^2(kq)\tilde{\theta}^T(kq)\hat{\varphi}(kq)\hat{\varphi}^T(kq)R^{-1}(kq)\hat{\varphi}(kq)e(kq) + \\
& 2\rho^3(kq)(\hat{\varphi}^T(kq)R^{-1}(kq)\hat{\varphi}(kq))^2e^2(kq) = \\
& \rho(kq)(\tilde{\theta}^T(kq)\hat{\varphi}(kq))^2 + \\
& 2\rho^3(kq)(\hat{\varphi}^T(kq)R^{-1}(kq)\hat{\varphi}(kq))^2e^2(kq) - \\
& 2\rho^2(kq)\tilde{\theta}^T(kq)\hat{\varphi}(kq)\hat{\varphi}^T(kq)R^{-1}(kq)\hat{\varphi}(kq) \times \\
& \omega^{-1}(kq)(\eta(kq) - v(kq)) - 2\rho^2(kq) \times \\
& \tilde{\theta}^T(kq)\hat{\varphi}(kq)\hat{\varphi}^T(kq)R^{-1}(kq)\hat{\varphi}(kq)\omega^{-1}(kq)v(kq) = \\
& \rho(kq)(\tilde{\theta}^T(kq)\hat{\varphi}(kq))^2 + \\
& 2\rho^3(kq)(\hat{\varphi}^T(kq)R^{-1}(kq)\hat{\varphi}(kq))^2e^2(kq) - \\
& 2\rho^2(kq)\tilde{\theta}^T(kq)\hat{\varphi}(kq)\hat{\varphi}^T(kq)R^{-1}(kq) \times \\
& \hat{\varphi}(kq)\omega^{-1}(kq)(\eta(kq) - v(kq)) - 2\rho^2(kq)\hat{\varphi}^T(kq) \times \\
& R^{-1}(kq)\hat{\varphi}(kq)\tilde{\theta}^T(kq - q)\hat{\varphi}(kq)\omega^{-1}(kq)v(kq) - \\
& 2\rho^3(kq)(\hat{\varphi}^T(kq)R^{-1}(kq)\hat{\varphi}(kq))^2\omega^{-1}(kq) \times \\
& (e(kq) - v(kq))v(kq) - 2\rho^3(kq) \times \\
& (\hat{\varphi}^T(kq)R^{-1}(kq)\hat{\varphi}(kq))^2\omega^{-1}(kq)v^2(kq). \quad (8)
\end{aligned}$$

定义

$$\begin{aligned}
H(kq) &= \hat{\varphi}^T(kq)R^{-1}(kq)\hat{\varphi}(kq) \geq 0, \\
\tilde{y}(kq) &= -\tilde{\theta}^T(kq)\hat{\varphi}(kq),
\end{aligned}$$

由式(6)~(8)可得

$$\begin{aligned}
V(kq) &\leq \\
V(kq - q) &+ \rho(kq)\tilde{y}^2(kq) + \\
2\rho^2(kq)H(kq)\tilde{y}(kq)\omega^{-1}(kq)(\eta(kq) - \\
v(kq)) &+ \rho^3(kq)H^2(kq)e^2(kq) - \\
2\rho^2(kq)H(kq)\tilde{\theta}^T(kq - q)\hat{\varphi}(kq)\omega^{-1}(kq)v(kq) - \\
2\rho^3(kq)H^2(kq)\omega^{-1}(kq)(e(kq) - v(kq))v(kq) - \\
2\rho^3(kq)H^2(kq)\omega^{-1}(kq)v^2(kq) - \\
2\rho(kq)\tilde{y}(kq)\omega^{-1}(kq)(\eta(kq) - v(kq)) + \\
2\rho(kq)\tilde{\theta}^T(kq - q)\hat{\varphi}(kq)\omega^{-1}(kq)v(kq) + \\
2\rho^2(kq)H(kq)\omega^{-1}(kq)(e(kq) - v(kq))v(kq) + \\
2\rho^2(kq)H(kq)\omega^{-1}(kq)v^2(kq) - \\
\rho^2(kq)H(kq)e^2(kq) = \\
V(kq - q) - [2\rho(kq)\tilde{y}(kq)\omega^{-1}(kq)(\eta(kq) - \\
v(kq)) - \rho(kq)\tilde{y}^2(kq) - 2\rho^2(kq) \times \\
H(kq)\tilde{y}(kq)\omega^{-1}(kq)(\eta(kq) - \\
v(kq))] + [2\rho^2(kq)H(kq)\omega^{-1}(kq)v^2(kq) - \\
2\rho^3(kq)H^2(kq)\omega^{-1}(kq)v^2(kq)] + \\
[\rho^3(kq)H^2(kq)e^2(kq) - \rho^2(kq)H(kq)e^2(kq)] + \\
[2\rho(kq)\tilde{\theta}^T(kq - q)\hat{\varphi}(kq)\omega^{-1}(kq)v(kq) + \\
2\rho^2(kq)H(kq)\omega^{-1}(kq)(e(kq) - v(kq))v(kq) - \\
2\rho^2(kq)H(kq)\tilde{\theta}^T(kq - q)\hat{\varphi}(kq)\omega^{-1}(kq)v(kq) - \\
2\rho^3(kq)H^2(kq)\omega^{-1}(kq)(e(kq) - v(kq))v(kq)]. \quad (9)
\end{aligned}$$

由于 $\rho(kq)$, $\tilde{\theta}^T(kq - q)\hat{\varphi}(kq)$, $\omega^{-1}(kq)$, $H(kq)$, $e(kq) - v(kq)$ 与 $v(kq)$ 不相关, 则对式(9)两端取条件期望可得到

$$\begin{aligned}
& E[V(kq) | F_{kq-q}] \leq \\
& V(kq - q) - E[2\rho(kq)\tilde{y}(kq)\omega^{-1}(kq)(\eta(kq) - \\
v(kq)) - \rho(kq)\tilde{y}^2(kq) - 2\rho^2(kq) \times \\
H(kq)\tilde{y}(kq)\omega^{-1}(kq)(\eta(kq) - v(kq)) | F_{kq-q}] - \\
& E[\rho^2(kq)H(kq)e^2(kq) - \\
\rho^3(kq)H^2(kq)e^2(kq) | F_{kq-q}] + \\
& E[2\rho^2(kq)H(kq)\omega^{-1}(kq) - \\
2\rho^3(kq)H^2(kq)\omega^{-1}(kq) | F_{kq-q}] \sigma_v^2, \quad (10)
\end{aligned}$$

其中

$$\begin{aligned}
& \rho^2(kq)H(kq)e^2(kq) - \rho^3(kq)H^2(kq)e^2(kq) = \\
& \rho^2(kq)H(kq)e^2(kq)(1 - \rho(kq)H(kq)) = \\
& \rho^2(kq)H(kq)e^2(kq)(1 - \\
& \rho(kq)\hat{\varphi}^T(kq)R^{-1}(kq)\hat{\varphi}(kq)).
\end{aligned}$$

由式(3)可知

$$R(kq) \geq \rho(kq)\hat{\varphi}^T(kq)\hat{\varphi}(kq),$$

则有

$$I \geq \rho(kq)R^{-1}(kq)\hat{\varphi}^T(kq)\hat{\varphi}(kq),$$

取行列式可得

$$1 \geq \rho(kq)\hat{\varphi}^T(kq)R^{-1}(kq)\hat{\varphi}(kq),$$

则有

$$\rho^2(kq)H(kq)e^2(kq) - \rho^3(kq)H^2(kq)e^2(kq) \geq 0. \quad (11)$$

同时, 式 (10) 中

$$\begin{aligned} & 2\rho^2(kq)H(kq)\omega^{-1}(kq) - \\ & 2\rho^3(kq)H^2(kq)\omega^{-1}(kq) = \\ & 2\rho^2(kq)H(kq)\omega^{-1}(kq)(1 - \rho(kq)H(kq)) = \\ & 2\rho^2(kq)H(kq)\omega^{-1}(kq)(1 - \\ & \rho(kq)\hat{\varphi}^T(kq)R^{-1}(kq)\hat{\varphi}(kq)) = \\ & 2\rho^2(kq)H(kq) \geq 0, \end{aligned} \quad (12)$$

$$\begin{aligned} & A(z)(\eta(kq) - v(kq)) = \\ & A(z)(y(kq) - \hat{\varphi}^T(kq)\hat{\theta}(kq) - v(kq)) = \\ & A(z)(y(kq) - \hat{\varphi}^T(kq)\hat{\theta}(kq) - \\ & v(kq) - y(kq) + \varphi^T(kq)\theta) = \\ & -A(z)\hat{\varphi}^T(kq)\hat{\theta}(kq) + A(z)\varphi^T(kq)\theta = \\ & -A(z)\hat{\varphi}^T(kq)\hat{\theta}(kq) + B(z)u(kq) = \\ & -\hat{\varphi}^T(kq)\hat{\theta}(kq) + \hat{\varphi}^T(kq)\theta = \\ & -\hat{\varphi}^T(kq)\tilde{\theta}(kq) = \tilde{y}(kq), \end{aligned}$$

则式 (10) 中

$$\begin{aligned} & 2\rho(kq)\tilde{y}(kq)\omega^{-1}(kq)(\eta(kq) - v(kq)) - \rho(kq)\tilde{y}^2(kq) - \\ & 2\rho^2(kq)H(kq)\tilde{y}(kq)\omega^{-1}(kq)(\eta(kq) - v(kq)) = \\ & 2\rho(kq)\tilde{y}(kq)(\eta(kq) - v(kq))(1 - \\ & \rho(kq)H(kq)\omega^{-1}(kq) - \rho(kq)\tilde{y}^2(kq) = \\ & 2\rho(kq)\tilde{y}(kq)(\eta(kq) - v(kq)) - \rho(kq)\tilde{y}^2(kq) = \\ & \rho(kq)\tilde{y}(kq)[2(\eta(kq) - v(kq)) - \\ & A(z)(\eta(kq) - v(kq))] = \\ & \rho(kq)\tilde{y}(kq)\frac{2 - A(z)}{A(z)}\tilde{y}(kq). \end{aligned}$$

由 A4 可知

$$\begin{aligned} & 2\rho(kq)\tilde{y}(kq)\omega^{-1}(kq)(\eta(kq) - v(kq)) - \rho(kq)\tilde{y}^2(kq) - \\ & 2\rho^2(kq)H(kq)\tilde{y}(kq)\omega^{-1}(kq)(\eta(kq) - v(kq)) \geq 0. \end{aligned} \quad (13)$$

定义

$$S(kq) = \sum_{i=1}^k [2\rho(iq)\tilde{y}(iq)\omega^{-1}(iq)(\eta(iq) -$$

$$\begin{aligned} & v(iq)) - \rho(iq)\tilde{y}^2(iq) - 2\rho^2(iq) \times \\ & H(iq)\tilde{y}(iq)\omega^{-1}(iq)(\eta(iq) - v(iq))] = \\ & \sum_{i=1}^k \rho(iq)\tilde{y}(iq)\frac{2 - A(z)}{A(z)}\tilde{y}(iq) \geq 0. \end{aligned}$$

并令 $Z(kq) = V(kq) + S(kq)$, 则式 (10) 可改写为

$$\begin{aligned} & E[Z(kq) | F_{kq-q}] \leq \\ & Z(kq - q) - E[\rho^2(kq)H(kq)e^2(kq) - \\ & \rho^3(kq)H^2(kq)e^2(kq) | F_{kq-q}] + \\ & 2\sigma_v^2 E[\rho^2(kq)H(kq)\omega^{-1}(kq) - \\ & \rho^3(kq)H^2(kq)\omega^{-1}(kq) | F_{kq-q}], \end{aligned} \quad (14)$$

其中

$$\begin{aligned} & \rho^2(kq)H(kq)\omega^{-1}(kq) - \rho^3(kq)H^2(kq)\omega^{-1}(kq) = \\ & \rho^2(kq)H(kq)\omega^{-1}(kq)(1 - \rho(kq)H(kq)) = \\ & \rho^2(kq)\hat{\varphi}^T(kq)R^{-1}(kq)\hat{\varphi}(kq). \end{aligned}$$

由引理 2、A1 可知

$$\lim_{k \rightarrow \infty} \sum_{i=1}^k \rho^2(iq)\hat{\varphi}^T(iq)R^{-1}(iq)\hat{\varphi}(iq) < \infty, \quad (15)$$

则根据引理 1, 由式 (11) ~ (13)、(15) 可知 (14) 中满足

$$\begin{cases} Z(kq) = V(kq) + S(kq) \xrightarrow{k \rightarrow \infty} Z_0 < \infty, \\ \sum_{i=1}^{\infty} \rho^2(iq)H(iq)e^2(iq) - \rho^3(iq)H^2(iq)e^2(iq) < \infty, \end{cases} \quad (16)$$

其中 Z_0 是有界随机变量. 对应地, 有

$$\lim_{k \rightarrow \infty} V(kq) = \lim_{k \rightarrow \infty} \tilde{\theta}^T(kq)R(kq)\tilde{\theta}(kq) < \infty, \quad (17)$$

$$\begin{aligned} & \lim_{k \rightarrow \infty} S(kq) = \\ & \lim_{k \rightarrow \infty} \sum_{i=1}^k \rho(iq)\tilde{y}(iq)\frac{2 - A(z)}{A(z)}\tilde{y}(iq) < \infty. \end{aligned} \quad (18)$$

由式 (17) 和引理 2 可知, $\|\tilde{\theta}(kq)\|^2$ 一致有界收敛.

综上, 定理 1 得证. \square

定理 2 在满足定理 1 和 A6 的条件下, 参数估计收敛到真值, 即 $\|\tilde{\theta}(kq)\|^2 \xrightarrow{k \rightarrow \infty} 0$.

证明 由式 (18) 可知

$$\sum_{i=1}^{\infty} \rho(iq)\tilde{y}^2(iq) < \infty,$$

根据 A1, 这里一定存在 N_1 , 当 $i \geq N_1$ 时满足

$$\begin{aligned} & 1 > \tilde{y}^2(iq) = \\ & \tilde{\theta}^T(iq)\hat{\varphi}(iq)\hat{\varphi}^T(iq)\tilde{\theta}(iq) = \\ & O(\rho(iq)) \rightarrow 0. \end{aligned}$$

由定理 1、引理 2、A5、A6 可知, $\tilde{\theta}(iq)$ 一致收敛于有界值, 令 $\tilde{\theta}(iq) \rightarrow \tilde{\theta}_\infty$, 当 N_1 充分大时有

$$\sum_{j=0}^{L-1} \tilde{y}^2(iq + jq) = \tilde{\theta}_{\infty}^T \sum_{j=0}^{L-1} \hat{\varphi}(iq + jq) \hat{\varphi}^T(iq + jq) \tilde{\theta}_{\infty} \rightarrow 0. \quad (19)$$

又由A6可知

$$0 \leq \alpha L \tilde{\theta}_{\infty}^T \tilde{\theta}_{\infty} \leq \tilde{\theta}_{\infty}^T \sum_{j=0}^{L-1} \hat{\varphi}(iq + jq) \hat{\varphi}^T(iq + jq) \tilde{\theta}_{\infty} \times \tilde{\theta}_{\infty}^T \sum_{j=0}^{L-1} \hat{\varphi}(iq + jq) \hat{\varphi}^T(iq + jq) \tilde{\theta}_{\infty} \leq \beta L \tilde{\theta}_{\infty}^T \tilde{\theta}_{\infty}. \quad (20)$$

由式(19)和(20)可知, 只有 $\tilde{\theta}_{\infty} = 0$ 才能保证两式结果一致, 由此说明参数估计收敛到了真值. \square

综上所述, 在充足的时间内, 该算法总能保证参数估计收敛到真值. 上述收敛性证明并没有对信息估计有过多要求(如A5、A6), 假设条件只是满足基本的有界条件, 其中的常参数也并非确数. 非线性模型依赖于实测数据建立, 其估计输出也必然有界, 在出现发散的情况下, 非线性模型总能够提供满足假设条件的信息估计, 确保递推进行下去, 直到参数收敛.

4 仿真实例

双率系统输出误差模型如下:

$$\begin{aligned} A(z)y(k) &= B(z)u(k) + v(k); \\ A(z) &= 1 + 0.412z^{-1} + 0.309z^{-2}, \\ B(z) &= 0.6804z^{-1} + 0.6303z^{-2}, \\ \theta &= [0.412, 0.309, 0.6804, 0.6303]^T. \end{aligned}$$

输入信号采用均值为0、方差为1的随机信号, 噪声采用均值为0、方差为0.01的白噪声. 采样周期比 $q = 50$, 即输出采样周期是输入采样周期的50倍. 为了说明切换机制的有效性, 增加基于常规辅助模型的改进随机牛顿递推算法作比较. 切换机制中, 次辅助模型采用神经网络模型, 并利用慢采样数据样本训练网络, 阈值设定为输出最大幅值的10倍. 利用式(4)辨识模型参数, 递推过程参数设定为: $\hat{\theta}(0) = \varepsilon$, $R(0) = a^2 \mathbf{I}$, 其中 ε 为充分小的实向量, a 为充分大的实数; 收敛因子 $\rho(kq) = 1/k$; 式(4)中 $\lambda = 10^{-2}$. 本文利用相对误差 $\delta \equiv \frac{\|\hat{\theta}(kq) - \theta\|}{\|\theta\|}$ 反映参数收敛情况, 如图2所示. 图3反映的是次辅助模型的动作, 规定若次辅助模型提供估计信息, 则认为次辅助模型动作为1, 否则为0.

从图3可以看出, 在 $t = 300$ 处或者第6个慢采样周期, 次辅助模型发生动作, 说明该处主辅助模型的中间不可测时刻估计发散, 次辅助模型及时提供信息估计. 根据图2, 在该动作之前两种算法的参数估计

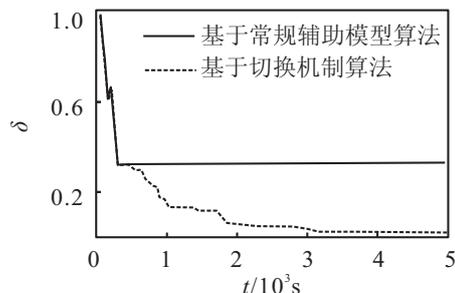


图2 相对误差曲线

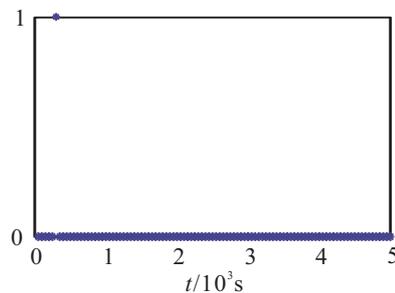


图3 次辅助模型动作图

的相对误差保持一致, 当出现发散现象时, 基于常规辅助模型的递推辨识算法无法进行, 因此其参数估计将会一直保持, 相对误差保持不变. 而本文提出的切换机制及时为递推算法提供信息估计, 确保算法进行, 表明了该切换机制的有效性. 另外, 可以看出次辅助模型动作次数较少, 尤其在后面主要依靠主辅助模型为递推算法提供信息估计, 即使这样, 次辅助模型对辨识过程仍至关重要, 若没有次辅助模型, 则在第6个慢采样周期之后递推无法进行.

5 结论

本文提出了基于双辅助模型的改进随机牛顿递推算法辨识双率系统模型, 并给出了参数估计一致收敛到真值的证明. 在仿真实例中可以看出, 本文给出的递推算法是有效的. 本文所提出方法的优势在于: 1) 增加非线性模型作为次辅助模型, 当中间不可测时刻输出估计发散时, 及时切换并给出下步递推所需信息估计, 确保递推进行以及参数估计最终收敛; 2) 在Hessian阵的递推中增加对称正定矩阵, 避免了当输入不充分或者不广泛时Hessian阵奇异或者接近奇异的情况, 确保递推进行. 然而本文提出的方法同样存在不足: 首先, 基于辅助模型的方法本身其收敛速度较慢, 即使最终收敛, 一般仍需较长时间, 而在实际应用中往往要求较短时间内参数收敛; 其次, 本文更偏向于在理论上说明收敛因子的存在性, 因此没有给出具体的表达式, 而收敛因子与收敛速度和收敛结果的要求相关, 今后将进一步研究收敛因子对该算法的影响并给出相应表达.

参考文献(References)

- [1] 倪博溢, 萧德云. 多采样率系统的辨识问题综述[J]. 控制理论与应用, 2009, 26(1): 62-68.
(Ni B Y, Xiao D Y. A survey on identification of multirate sampled systems[J]. Control Theory & Applications, 2009, 26(1): 62-68.)
- [2] Cristian G, Christian L, Marco M, et al. Multi-rate optimizing control of simulated moving beds[J]. J of Process Control, 2010, 20(4): 490-505.
- [3] Rossiter J A, Sheng T, Chen S L, et al. Interpretations of and options in dual-rate predictive control[J]. J of Process Control, 2005, 15(2): 135-148.
- [4] Chen T. On stability robustness of a dual-rate control system[J]. IEEE Trans on Automatic Control, 1994, 39(1): 164-167.
- [5] Li W, Hang Z, Shah S L. Subspace identification for FDI in systems with non-uniformly sampled multirate data[J]. Automatica, 2006, 42(4): 619-627.
- [6] Han L L, Ding F. Identification for multirate multi-input systems using the multi-innovation identification theory[J]. Computers and Mathematics with Applications, 2009, 57(9): 1438-1449.
- [7] Qin P, Kanae S, Yang Z, et al. Identification of lifted models for general dual-rate sampled data systems using N5SID algorithm[J]. IEEE Trans on Electronics, Information and Systems, 2008, 128(5): 788-794.
- [8] 吕立华, 宋执环, 李平. 一种基于小波多分辨分析的多率采样系统辨识方法[J]. 控制理论与应用, 2002, 19(2): 225-234.
(Lv L H, Song Z H, Li P. An identification method for multirate sampled-data systems based on wavelet multiresolution analysis[J]. Control Theory & Applications, 2002, 19(2): 225-234.)
- [9] Lu W P, Fisher G. Output estimation with multi-rate sampling[J]. Int J of Control, 1988, 48(1): 149-160.
- [10] Lu W P, Fisher G. Least-squares output estimation with multirate sampling[J]. IEEE Trans on Automatic Control, 1989, 34(6): 669-672.
- [11] Ding F, Chen T. Hierarchical identification of lifted state-space models for general dual-rate systems[J]. IEEE Trans on Circuits and Systems I: Regular Papers, 2005, 52(6): 1179-1187.
- [12] 丁锋, 萧德云. 多变量系统状态空间模型的递阶辨识[J]. 控制与决策, 2005, 20(8): 848-853.
(Ding F, Xiao D Y. Hierarchical identification of state space models for multivariable systems[J]. Control and Decision, 2005, 20(8): 848-853.)
- [13] Raghavan H, Gopaluni B, Shah S L, et al. Gray-box identification of dynamic models for the bleaching operation in a pulp mill[J]. J of Process Control, 2005, 15(4): 451-468.
- [14] Ding F, Chen T. Combined parameter and output estimation of dual-rate systems using an auxiliary model[J]. Automatica, 2004, 40(10): 1739-1748.
- [15] Ding F, Chen T. Modeling and identification of multirate systems[J]. Acta Automatica Sinica, 2005, 31(1): 105-122.
- [16] Liu Y J, Xiao Y S, Zhao X L. Multi-innovation stochastic gradient algorithm for multiple-input single-output systems using the auxiliary model[J]. Applied Mathematics and Computation, 2009, 215(4): 1477-1483.
- [17] Ferreira O P, Goncalves M L N, Oliveira P R. Local convergence analysis of inexact Gauss-Newton like methods under majorant condition[J]. J of Computational and Applied Mathematics, 2012, 236(9): 2487-2498.
- [18] Ljung L, Soderstrom T. Theory and practice of recursive identification[M]. Cambridge: MIT Press, 1983: 361-367.
- [19] Khargonekar P P, Poolla K, Tannenbaum A. Robust control of linear time-invariant plants using periodic compensation[J]. IEEE Trans on Automatic Control, 1985, 30(11): 1088-1096.
- [20] 冯瑞, 张玥杰, 张艳珠, 等. 基于加权支持向量机的移动建模方法及其在软测量中的应用[J]. 自动化学报, 2004, 30(3): 436-441.
(Zhang R, Zhang Y J, Zhang Y Z, et al. Drifting modeling method using weighted support vector machines with application to soft sensor[J]. Acta Automatica Sinica, 2004, 30(3): 436-441.)
- [21] Wang Y H, Huang D X, Gao D J, et al. Wavelet networks based soft sensor and predictive control in fermentation process[J]. Computer Aided Chemical Engineering, 2003, 15: 1222-1227.
- [22] Goodwin G C, Sin K S. Adaptive filtering prediction and control[M]. Englewood Cliffs: Prentice-Hall, 1984: 551-560.
- [23] Wang D Q, Ding F. Performance analysis of the auxiliary models based multi-innovation stochastic gradient estimation algorithm for output error systems[J]. Digital Signal Processing, 2010, 20(3): 750-762.

(责任编辑: 孙艺红)