



中国科技期刊卓越行动计划项目入选期刊

# 控制与决策

CONTROL AND DECISION



## 基于数据滤波的随机梯度辨识方法

丁锋, 郑嘉芸, 张霄, 徐玲

引用本文:

丁锋, 郑嘉芸, 张霄, 徐玲. 基于数据滤波的随机梯度辨识方法[J]. 控制与决策, 2024, 39(7): 2259–2266.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2022.1867>

## 您可能感兴趣的其他文章

### Articles you may be interested in

#### 含有分数阶有色关联噪声的分数阶系统的卡尔曼滤波器设计

Design of Kalman filter for fractional-order systems with correlated fractional-order colored noises

控制与决策. 2021, 36(7): 1672–1678 <https://doi.org/10.13195/j.kzyjc.2019.1418>

#### 基于KRLS的非均匀采样非线性系统辨识

Identification of non-uniformly sampled nonlinear systems based on KRLS

控制与决策. 2021, 36(12): 3049–3055 <https://doi.org/10.13195/j.kzyjc.2020.0962>

#### 基于向量集逆区间滤波的故障观测器设计

Vector set inversion interval filtering based fault observer design

控制与决策. 2021, 36(12): 2973–2981 <https://doi.org/10.13195/j.kzyjc.2020.0212>

#### 基于转换量测容积卡尔曼滤波器带多普勒量测的目标跟踪算法

Converted measurement cubature Kalman filter for Doppler-assisted target tracking

控制与决策. 2021, 36(6): 1425–1434 <https://doi.org/10.13195/j.kzyjc.2019.1070>

#### 基于变速趋近律的Buck型变换器抗扰动控制

Disturbance rejection control of Buck converters based on variable rate reaching law

控制与决策. 2021, 36(4): 893–900 <https://doi.org/10.13195/j.kzyjc.2019.1073>

# 基于数据滤波的随机梯度辨识方法

丁锋<sup>†</sup>, 郑嘉芸, 张霄, 徐玲

(江南大学 物联网工程学院, 江苏 无锡 214122)

**摘要:** 针对有色噪声干扰下的随机系统, 利用数据滤波技术, 对输入输出数据进行滤波, 将具有滑动平均噪声的原始系统转换为白噪声干扰下的系统, 提出有限脉冲响应滑动平均系统的滤波增广随机梯度算法, 并对该算法进行收敛性分析. 此外, 为了提高参数估计的精度和加快算法的收敛速度, 使用多新息辨识理论提出滤波多新息增广随机梯度算法, 并分析其收敛性. 与增广随机梯度算法相比, 所提出的滤波增广随机梯度算法和滤波多新息增广随机梯度算法可以得到更高精度的参数估计. 最后, 通过仿真实例表明了所提出算法的有效性.

**关键词:** 参数估计; 多新息辨识; 数据滤波; 随机系统; 随机梯度; 收敛性分析

中图分类号: TP273

文献标志码: A

DOI: 10.13195/j.kzyjc.2022.1867

引用格式: 丁锋, 郑嘉芸, 张霄, 等. 基于数据滤波的随机梯度辨识方法[J]. 控制与决策, 2024, 39(7): 2259-2266.

## Filtering-based stochastic gradient identification methods

DING Feng<sup>†</sup>, ZHENG Jia-yun, ZHANG Xiao, XU Ling

(School of Internet of Things Engineering, Jiangnan University, Wuxi 214122, China)

**Abstract:** This paper studies the parameter identification of stochastic systems with colored noises. Using the data filtering technology to filter the input and output data, which converts the original system with moving average noise to the system with white noise, we propose the filtering-based extended stochastic gradient algorithm and analyze its convergence. In addition, in order to improve the parameter estimation accuracy and accelerate the convergence rate, a filtering-based multi-innovation extended stochastic gradient algorithm is proposed by using the multi-innovation identification theory and its convergence is analyzed. Compared with the extended stochastic gradient algorithm, the proposed filtering-based extended stochastic gradient algorithm and the filtering-based multi-innovation extended stochastic gradient algorithm can obtain higher precision parameter estimates. Finally, the simulation results indicate that the proposed algorithms are effective.

**Keywords:** parameter estimation; multi-innovation identification; data filtering; stochastic systems; stochastic gradient; convergence analysis

## 0 引言

系统辨识广泛应用于许多科学和工程领域, 如航空航天、生物医学、自适应控制<sup>[1]</sup>和故障诊断<sup>[2-3]</sup>等都需要用到辨识技术, 根据观测数据来确定数学模型. 数学模型是控制器设计<sup>[4]</sup>和系统分析<sup>[5]</sup>的基础, 对于有限脉冲响应(FIR)系统, 目前已经提出了许多参数估计方法. 例如, Wang等<sup>[6]</sup>针对FIR系统研究了自适应控制以及辨识问题. Munker等<sup>[7]</sup>研究了FIR模型的非线性系统辨识.

最小二乘算法和随机梯度算法是辨识领域中两种常用的方法. 最小二乘算法具有较快的收敛速度, 但是该算法需要计算协方差矩阵, 计算量较大. 而随

机梯度算法的计算量较小<sup>[8]</sup>, 但其收敛速度相对较低. 为改进随机梯度辨识方法的收敛速度, 可通过扩展新息维度推导出多新息随机梯度算法<sup>[9]</sup>. Fan等<sup>[10]</sup>便是在随机梯度算法的基础上提出多新息随机梯度算法估计分段线性输入非线性FIR-MA系统的参数. Ma等<sup>[11]</sup>提出了基于数据滤波的多新息最小二乘算法, 在减小计算量的同时提升了收敛速度.

滤波方法是一类新的参数估计方法. 辨识领域的滤波技术不同于通信和信号处理领域的滤波器, 后者的作用是去除噪声信号中的干扰, 前者是在系统方程两边同乘以滤波器将受有色噪声干扰的系统转换为受白噪声影响的系统, 以提高参数识别精度<sup>[12]</sup>. 近

收稿日期: 2022-10-28; 录用日期: 2023-04-13.

基金项目: 国家自然科学基金项目(62273167).

责任编委: 柴利.

<sup>†</sup>通讯作者. E-mail: fding@jiangnan.edu.cn.

年来出现了许多基于滤波技术的辨识方法,文献[13]针对ARX模型的参数估计提出了卡尔曼滤波辨识方法.文献[14]通过使用数据滤波技术推导了双线性状态空间系统的增广梯度迭代算法,提高了参数估计的精度.

本文研究有限脉冲响应滑动平均(FIR-MA)系统的参数辨识问题,推导增广随机梯度(ESG)算法,通过使用数据滤波技术提出滤波增广随机梯度(F-ESG)算法,并分析该算法的收敛性.为了加快参数的收敛速度和得到更精确的参数估计,引入新息长度,将新息标量扩展为新息向量,提出基于滤波的多新息增广随机梯度(F-MI-ESG)算法,并给出算法的收敛性证明.最后通过两个仿真例子表明了所提出算法的有效性,给出一些结论.

## 1 辨识模型

考虑如下有限脉冲响应滑动平均(FIR-MA)模型描述的随机控制系统:

$$y(t) = B(z)u(t) + D(z)v(t). \quad (1)$$

其中:  $\{u(t)\}$  和  $\{y(t)\}$  分别为系统的输入和输出序列,  $\{v(t)\}$  为零均值不可测白噪声序列,  $B(z)$  和  $D(z)$  为后移算子  $z^{-1}$  的多项式 ( $z^{-1}y(t) = y(t-1)$  或  $zy(t) = y(t+1)$ ), 定义为

$$B(z) := b_1z^{-1} + b_2z^{-2} + \dots + b_{n_b}z^{-n_b},$$

$$D(z) := 1 + d_1z^{-1} + d_2z^{-2} + \dots + d_{n_d}z^{-n_d}.$$

定义干扰噪声  $w(t) := D(z)v(t)$ , 是一个MA噪声, 且是有色噪声. 对辨识而言, 这里要求噪声模型  $D(z)$  (即滤波器) 是稳定的, 这样滤波后的变量才不至于发散. 假设阶次  $n_b$  和  $n_d$  已知, 记  $n := n_b + n_d$ . 定义如下参数向量和信息向量:

$$\mathbf{b} := [b_1, b_2, \dots, b_{n_b}]^T \in \mathbb{R}^{n_b},$$

$$\mathbf{d} := [d_1, d_2, \dots, d_{n_d}]^T \in \mathbb{R}^{n_d},$$

$$\boldsymbol{\vartheta} := [\mathbf{b}^T, \mathbf{d}^T]^T \in \mathbb{R}^n,$$

$$\boldsymbol{\varphi}(t) := [u(t-1), u(t-2), \dots, u(t-n_b),$$

$$v(t-1), v(t-2), \dots, v(t-n_d)]^T \in \mathbb{R}^n.$$

由式(1)可得

$$y(t) = \boldsymbol{\varphi}^T(t)\boldsymbol{\vartheta} + v(t). \quad (2)$$

式(2)是FIR-MA系统(1)的辨识模型.

取滤波器  $1/D(z)$ , 定义系统的滤波输入变量  $u_f(t)$  和系统的滤波输出变量  $y_f(t)$  分别为

$$u_f(t) := \frac{1}{D(z)}u(t), \quad (3)$$

$$y_f(t) := \frac{1}{D(z)}y(t). \quad (4)$$

定义滤波信息向量

$$\boldsymbol{\varphi}_f(t) := [u_f(t-1), u_f(t-2), \dots, u_f(t-n_b),$$

$$y_f(t-1), \dots, y_f(t-n_d)]^T \in \mathbb{R}^n. \quad (5)$$

式(1)两边同除以  $D(z)$ , 可得

$$y_f(t) = B(z)u_f(t) + v(t). \quad (6)$$

由式(3)和(4)得

$$u(t) = D(z)u_f(t), \quad y(t) = D(z)y_f(t).$$

进一步可得

$$u_f(t) = [1 - D(z)]u_f(t) + u(t) =$$

$$u(t) - d_1u_f(t-1) - d_2u_f(t-2) -$$

$$\dots - d_{n_d}u_f(t-n_d), \quad (7)$$

$$y_f(t) = [1 - D(z)]y_f(t) + y(t) =$$

$$y(t) - d_1y_f(t-1) - d_2y_f(t-2) -$$

$$\dots - d_{n_d}y_f(t-n_d). \quad (8)$$

将式(8)代入(6), 得到

$$[1 - D(z)]y_f(t) + y(t) = B(z)u_f(t) + v(t),$$

即

$$y(t) = [D(z) - 1]y_f(t) + B(z)u_f(t) + v(t) =$$

$$[u_f(t-1), u_f(t-2), \dots, u_f(t-n_b)]\mathbf{b} +$$

$$[y_f(t-1), y_f(t-2), \dots, y_f(t-n_d)]\mathbf{d} + v(t) =$$

$$\boldsymbol{\varphi}_f^T(t)\boldsymbol{\vartheta} + v(t). \quad (9)$$

式(9)是FIR-MA系统(1)的滤波辨识模型. 参数向量  $\boldsymbol{\vartheta}$  包含了所有参数  $b_i$  和  $d_i$ , 滤波信息向量由滤波输入  $u_f(t-i)$  和滤波输出  $y_f(t-i)$  组成. 本文的目的是使用数据滤波技术、梯度搜索法和多新息理论推导估计参数向量的新算法.

## 2 增广随机梯度算法

对于辨识模型(2), 定义梯度准则函数

$$J_1(\boldsymbol{\vartheta}) := [y(t) - \boldsymbol{\varphi}^T(t)\boldsymbol{\vartheta}]^2.$$

令  $\hat{\boldsymbol{\vartheta}}(t) := [\hat{\mathbf{b}}^T(t), \hat{\mathbf{d}}^T(t)]^T$  为参数向量  $\boldsymbol{\vartheta} = [\mathbf{b}^T, \mathbf{d}^T]^T$  在时刻  $t$  的估计,  $1/r(t)$  为收敛因子. 使用负梯度搜索极小化准则函数  $J_1(\boldsymbol{\vartheta})$ , 得到如下梯度递推关系:

$$\hat{\boldsymbol{\vartheta}}(t) = \hat{\boldsymbol{\vartheta}}(t-1) - \frac{1}{r(t)}\text{grad}[J_1(\hat{\boldsymbol{\vartheta}}(t-1))] =$$

$$\hat{\boldsymbol{\vartheta}}(t-1) + \frac{\boldsymbol{\varphi}(t)}{r(t)}[y(t) - \boldsymbol{\varphi}^T(t)\hat{\boldsymbol{\vartheta}}(t-1)];$$

$$r(t) = r(t-1) + \|\boldsymbol{\varphi}(t)\|^2, \quad r(0) = 1.$$

$\hat{\boldsymbol{\vartheta}}(t)$  的初值通常选择为  $\hat{\boldsymbol{\vartheta}}(0) = \mathbf{1}_n/p_0$ . 由于信息向量  $\boldsymbol{\varphi}(t)$  中包含了不可测噪声  $v(t-i)$ , 不可测噪声  $v(t)$  用其估计  $\hat{v}(t) := y(t) - \hat{\boldsymbol{\varphi}}^T(t)\hat{\boldsymbol{\vartheta}}(t)$  代替, 即可得到增广随机梯度 (ESG) 算法

$$\hat{\boldsymbol{\vartheta}}(t) = \hat{\boldsymbol{\vartheta}}(t-1) + \frac{\hat{\boldsymbol{\varphi}}(t)}{r(t)}[y(t) - \hat{\boldsymbol{\varphi}}^T(t)\hat{\boldsymbol{\vartheta}}(t-1)]; \quad (10)$$

$$r(t) = r(t-1) + \|\hat{\boldsymbol{\varphi}}(t)\|^2, \quad r(0) = 1; \quad (11)$$

$$\hat{\boldsymbol{\varphi}}(t) = [u(t-1), u(t-2), \dots, u(t-n_b), \hat{v}(t-1), \hat{v}(t-2), \dots, \hat{v}(t-n_d)]^T; \quad (12)$$

$$\hat{v}(t) = y(t) - \hat{\boldsymbol{\varphi}}^T(t)\hat{\boldsymbol{\vartheta}}(t); \quad (13)$$

$$\hat{\mathbf{b}}(t) = [\hat{b}_1(t), \hat{b}_2(t), \dots, \hat{b}_{n_b}(t)]^T; \quad (14)$$

$$\hat{\mathbf{d}}(t) = [\hat{d}_1(t), \hat{d}_2(t), \dots, \hat{d}_{n_d}(t)]^T; \quad (15)$$

$$\hat{\boldsymbol{\vartheta}}(t) = [\hat{\mathbf{b}}^T(t), \hat{\mathbf{d}}^T(t)]^T. \quad (16)$$

### 3 基于滤波的增广随机梯度算法

对于 FIR-MA 系统的滤波辨识模型 (9), 定义梯度准则函数

$$J_2(\boldsymbol{\vartheta}) := [y(t) - \boldsymbol{\varphi}_f^T(t)\boldsymbol{\vartheta}]^2.$$

使用负梯度搜索极小化准则函数  $J_2(\boldsymbol{\vartheta})$ , 可得到如下梯度递推关系:

$$\begin{aligned} \hat{\boldsymbol{\vartheta}}(t) &= \hat{\boldsymbol{\vartheta}}(t-1) - \frac{1}{r(t)} \text{grad}[J_2(\hat{\boldsymbol{\vartheta}}(t-1))] = \\ &\hat{\boldsymbol{\vartheta}}(t-1) + \frac{\boldsymbol{\varphi}_f(t)}{r(t)}[y(t) - \boldsymbol{\varphi}_f^T(t)\hat{\boldsymbol{\vartheta}}(t-1)]; \end{aligned} \quad (17)$$

$$r(t) = r(t-1) + \|\boldsymbol{\varphi}_f(t)\|^2, \quad r(0) = 1. \quad (18)$$

信息向量  $\boldsymbol{\varphi}_f(t)$  中包含了未知变量  $y_f(t-i)$  和  $u_f(t-i)$ , 解决方法是根据式 (5) 中  $\boldsymbol{\varphi}_f(t)$  的结构, 将其中的未知滤波变量  $y_f(t-i)$  和  $u_f(t-i)$  用其对应的估计  $\hat{y}_f(t-i)$  和  $\hat{u}_f(t-i)$  来代替, 滤波中使用的未知多项式  $D(z)$  也用其在时刻  $t$  的估计  $\hat{D}(t, z)$  代替. 定义  $\boldsymbol{\varphi}_f(t)$  的估计为

$$\hat{\boldsymbol{\varphi}}_f(t) := [\hat{y}_f(t-1), \hat{y}_f(t-2), \dots, \hat{y}_f(t-n_d), \hat{u}_f(t-1), \dots, \hat{u}_f(t-n_b)]^T \in \mathbb{R}^n. \quad (19)$$

用参数向量  $\mathbf{d}$  的估计  $\hat{\mathbf{d}}(t) := [\hat{d}_1(t), \hat{d}_2(t), \dots, \hat{d}_{n_d}(t)]^T \in \mathbb{R}^{n_d}$  构造多项式  $D(z)$  的估计, 有

$$\hat{D}(t, z) = 1 + \hat{d}_1(t)z^{-1} + \hat{d}_2(t)z^{-2} + \dots + \hat{d}_{n_d}(t)z^{-n_d}.$$

由式 (7) 和 (8) 可知  $u_f(t)$  和  $y_f(t)$  的估计分别为

$$\begin{aligned} \hat{u}_f(t) &= [1 - \hat{D}(t, z)]\hat{u}_f(t) + u(t) = \\ &u(t) - \hat{d}_1(t)\hat{u}_f(t-1) - \hat{d}_2(t)\hat{u}_f(t-2) - \\ &\dots - \hat{d}_{n_d}(t)\hat{u}_f(t-n_d), \end{aligned} \quad (20)$$

$$\begin{aligned} \hat{y}_f(t) &= [1 - \hat{D}(t, z)]\hat{y}_f(t) + y(t) = \\ &y(t) - \hat{d}_1(t)\hat{y}_f(t-1) - \hat{d}_2(t)\hat{y}_f(t-2) - \\ &\dots - \hat{d}_{n_d}(t)\hat{y}_f(t-n_d). \end{aligned} \quad (21)$$

将式 (17) 和 (18) 中  $\boldsymbol{\varphi}_f(t)$  用其估计  $\hat{\boldsymbol{\varphi}}_f(t)$  代替, 联立式 (19)~(21), 可以得到估计 FIR-MA 系统参数向量  $\boldsymbol{\vartheta}$  基于滤波的增广随机梯度 (F-ESG) 算法如下:

$$\begin{aligned} \hat{\boldsymbol{\vartheta}}(t) &= \hat{\boldsymbol{\vartheta}}(t-1) + \frac{\hat{\boldsymbol{\varphi}}_f(t)}{r(t)}[y(t) - \hat{\boldsymbol{\varphi}}_f^T(t)\hat{\boldsymbol{\vartheta}}(t-1)] = \\ &\hat{\boldsymbol{\vartheta}}(t-1) + \frac{\hat{\boldsymbol{\varphi}}_f(t)}{r(t)}e(t); \end{aligned} \quad (22)$$

$$e(t) = y(t) - \hat{\boldsymbol{\varphi}}_f^T(t)\hat{\boldsymbol{\vartheta}}(t-1); \quad (23)$$

$$r(t) = r(t-1) + \|\hat{\boldsymbol{\varphi}}_f(t)\|^2, \quad r(0) = 1; \quad (24)$$

$$\begin{aligned} \hat{\boldsymbol{\varphi}}_f(t) &= [\hat{u}_f(t-1), \hat{u}_f(t-2), \dots, \hat{u}_f(t-n_b), \\ &\hat{y}_f(t-1), \dots, \hat{y}_f(t-n_d)]^T; \end{aligned} \quad (25)$$

$$\begin{aligned} \hat{y}_f(t) &= y(t) - [\hat{y}_f(t-1), \hat{y}_f(t-2), \dots, \\ &\hat{y}_f(t-n_d)]\hat{\mathbf{d}}(t); \end{aligned} \quad (26)$$

$$\begin{aligned} \hat{u}_f(t) &= u(t) - [\hat{u}_f(t-1), \hat{u}_f(t-2), \dots, \\ &\hat{u}_f(t-n_d)]\hat{\mathbf{d}}(t); \end{aligned} \quad (27)$$

$$\hat{\mathbf{b}}(t) = [\hat{b}_1(t), \hat{b}_2(t), \dots, \hat{b}_{n_b}(t)]^T; \quad (28)$$

$$\hat{\mathbf{d}}(t) = [\hat{d}_1(t), \hat{d}_2(t), \dots, \hat{d}_{n_d}(t)]^T; \quad (29)$$

$$\hat{\boldsymbol{\vartheta}}(t) = [\hat{\mathbf{b}}^T(t), \hat{\mathbf{d}}^T(t)]^T. \quad (30)$$

F-ESG 算法 (22)~(30) 计算参数估计向量  $\hat{\boldsymbol{\vartheta}}(t)$  的步骤如下.

step 1: 令  $t = 1$ , 置初值  $\hat{\boldsymbol{\vartheta}}(0) = \mathbf{1}_n/p_0, r(0) = 1, \hat{y}_f(t-i) = 1/p_0, \hat{u}_f(t-i) = 1/p_0, i = 1, 2, \dots, \max[n_b, n_d], p_0 = 10^6$ . 给定数据长度  $L_e$ .

step 2: 采集输入和输出数据  $u(t)$  和  $y(t)$ .

step 3: 按照式 (25) 构造滤波信息向量  $\hat{\boldsymbol{\varphi}}_f(t)$ , 利用式 (23) 计算新息  $e(t)$ , 利用式 (24) 计算  $r(t)$ .

step 4: 利用式 (22) 更新参数估计向量  $\hat{\boldsymbol{\vartheta}}(t)$ .

step 5: 根据式 (30) 从  $\hat{\boldsymbol{\vartheta}}(t)$  中读取  $\hat{\mathbf{b}}(t)$  和  $\hat{\mathbf{d}}(t)$ .

step 6: 利用式 (26) 计算滤波输出  $\hat{y}_f(t)$ , 利用式 (27) 计算滤波输入  $\hat{u}_f(t)$ .

step 7: 如果  $t < L_e$ , 则  $t$  增加 1, 转至 step 2; 否则输出参数估计  $\hat{\boldsymbol{\vartheta}}(t)$ , 终止递推计算过程.

对于时不变参数估计算法, 要求算法的增益趋于零 (即收敛因子), 才能保证参数估计收敛性能.  $r(t)$  定义为信息向量范数的平方和, 它是随观测数据不断增加而趋于无穷, 从而可保证参数估计误差收敛于零.

在 F-ESG 辨识算法 (22)~(30) 中,  $e(t) = y(t) - \hat{\boldsymbol{\varphi}}_f^T(t)\hat{\boldsymbol{\vartheta}}(t-1) \in \mathbb{R}$  为辨识新息. 定义辨识残差

$$\hat{v}(t) := y(t) - \hat{\boldsymbol{\varphi}}_f^T(t)\hat{\boldsymbol{\vartheta}}(t) \in \mathbb{R}. \quad (31)$$

**引理1** 对于F-ESG辨识算法(22)~(30),新息与残差满足

$$\hat{v}(t) = \frac{r(t-1)}{r(t)}e(t).$$

**证明** 将式(22)代入(31),有

$$\begin{aligned} \hat{v}(t) &= y(t) - \hat{\varphi}_f^T(t) \left[ \hat{\vartheta}(t-1) + \frac{\hat{\varphi}_f(t)}{r(t)}e(t) \right] = \\ & y(t) - \hat{\varphi}_f^T(t)\hat{\vartheta}(t-1) - \frac{\|\hat{\varphi}_f(t)\|^2}{r(t)}e(t) = \\ & e(t) - \frac{\|\hat{\varphi}_f(t)\|^2}{r(t)}e(t) = \left[ 1 - \frac{\|\hat{\varphi}_f(t)\|^2}{r(t)} \right] e(t) = \\ & \frac{r(t-1)}{r(t)}e(t). \quad \square \end{aligned}$$

**引理2** 设非负序列  $\{x(t)\}$ 、 $\{\alpha_t\}$  和  $\{\beta_t\}$  满足如下关系<sup>[15]</sup>:

$$x(t) \leq (1 - \alpha_t)x(t-1) + \beta_t, \quad t \geq 0,$$

且  $\alpha_t \in [0, 1)$ ,  $\sum_{t=1}^{\infty} \alpha_t = \infty$ ,  $x(0) < \infty$ , 则有

$$\limsup_{t \rightarrow +\infty} x(t) \leq \limsup_{t \rightarrow +\infty} \frac{\beta_t}{\alpha_t},$$

其中假设有关极限存在.

**引理3** 对于F-ESG辨识算法(22)~(30),假设存在常数  $0 < \alpha \leq \beta < \infty$  和整数  $N \geq n$ , 有下列强持续激励条件(SPE)成立<sup>[15]</sup>:

$$\begin{aligned} \alpha \mathbf{I}_n &\leq \frac{1}{N} \sum_{i=0}^{N-1} \hat{\varphi}_f(t+i)\hat{\varphi}_f^T(t+i) \leq \beta \mathbf{I}_n, \quad \text{a.s.}, \\ &t > 0. \end{aligned}$$

则式(24)中  $r(t)$  满足

$$\begin{aligned} (t - N + 1)n\alpha &\leq r(t) \leq (t + N)n\beta + 1, \quad \text{a.s.}, \\ &t \geq N. \end{aligned}$$

**引理4** 对于F-ESG算法(22)~(30),定义转移矩阵

$$\begin{aligned} \Xi(t+1, i) &:= \left[ \mathbf{I} - \frac{\hat{\varphi}_f(t)\hat{\varphi}_f^T(t)}{r(t)} \right] \Xi(t, i), \\ \Xi(i, i) &= \mathbf{I}_n. \end{aligned}$$

引理3中SPE条件成立,则有

$$\begin{aligned} \rho_t &:= \\ \lambda_{\max}[\Xi^T(t+N, t)\Xi(t+N, t)] &\leq \\ 1 - \frac{\alpha^2}{n\beta\{[n(t+2N)\beta+1]^{1/2} + N\sqrt{nN\beta}\}^2}, &\text{ a.s.} \end{aligned}$$

其中:  $t > 0$ ,  $\lambda_{\max}[\mathbf{X}]$  为对称方阵  $\mathbf{X}$  的最大特征值.

**定理1** 对于F-ESG算法(22)~(30),假设SPE条件成立,系统输入  $u(t)$  和系统输出  $y(t)$  有上界,即

$|u(t)| \leq \sqrt{\beta}$ ,  $|y(t)| \leq \sqrt{\beta}$ , 那么参数估计  $\hat{\vartheta}(t)$  是有界收敛的.

在此辨识方法的收敛性分析中,通常不显示强调系统稳定,而是通过持续激励条件对观测输入输出数据有界来加以约束.

### 4 基于滤波的多新息增广随机梯度算法

随机梯度算法的收敛性较慢,因此借助于多新息辨识理论,新息是能改进参数估计精度的有用信息.多新息策略将标量新息扩展为新息向量,引入多新息不仅可以加快算法的收敛速度,而且能够提高算法的参数估计精度.通过扩展新息维数,推导出基于滤波的多新息增广随机梯度(F-MI-ESG)算法.设正整数  $p$  为新息长度,基于F-ESG算法(22)~(30),将系统滤波信息向量  $\hat{\varphi}_f(t)$  和输出  $y(t)$  扩展为堆积滤波信息矩阵  $\hat{\Phi}_f(p, t) \in \mathbb{R}^{n \times p}$  和堆积输出向量  $\mathbf{Y}(p, t) \in \mathbb{R}^p$ , 有

$$\hat{\Phi}_f(p, t) := [\hat{\varphi}_f(t), \dots, \hat{\varphi}_f(t-p+1)], \quad (32)$$

$$\mathbf{Y}(p, t) := [y(t), y(t-1), \dots, y(t-p+1)]^T. \quad (33)$$

将标量新息  $e(t) = y(t) - \hat{\varphi}_f^T(t)\hat{\vartheta}(t-1) \in \mathbb{R}$  扩展为新息向量

$$\begin{aligned} \mathbf{E}(p, t) &= \begin{bmatrix} y(t) - \hat{\varphi}_f^T(t)\hat{\vartheta}(t-1) \\ y(t-1) - \hat{\varphi}_f^T(t-1)\hat{\vartheta}(t-1) \\ \vdots \\ y(t-p+1) - \hat{\varphi}_f^T(t-p+1)\hat{\vartheta}(t-1) \end{bmatrix} = \\ &\mathbf{Y}(p, t) - \hat{\Phi}_f^T(p, t)\hat{\vartheta}(t-1) \in \mathbb{R}^p. \quad (34) \end{aligned}$$

用堆积滤波信息矩阵  $\hat{\Phi}_f(p, t)$  代替式(22)和(24)中的  $\hat{\varphi}_f(t)$ , 得到式(35)和(37).联立式(32)~(34)和(25)~(30)可以得到估计FIR-MA系统参数向量  $\vartheta$  基于滤波的多新息增广随机梯度(F-MI-ESG)算法

$$\hat{\vartheta}(t) = \hat{\vartheta}(t-1) + \frac{\hat{\Phi}_f(p, t)}{r(t)}\mathbf{E}(p, t); \quad (35)$$

$$\mathbf{E}(p, t) = \mathbf{Y}(p, t) - \hat{\Phi}_f^T(p, t)\hat{\vartheta}(t-1); \quad (36)$$

$$r(t) = r(t-1) + \|\hat{\Phi}_f(p, t)\|^2, \quad r(0) = 1; \quad (37)$$

$$\mathbf{Y}(p, t) = [y(t), y(t-1), \dots, y(t-p+1)]^T; \quad (38)$$

$$\hat{\Phi}_f(p, t) = [\hat{\varphi}_f(t), \dots, \hat{\varphi}_f(t-p+1)]; \quad (39)$$

$$\begin{aligned} \hat{\varphi}_f(t) &= [\hat{u}_f(t-1), \hat{u}_f(t-2), \dots, \hat{u}_f(t-n_b), \\ &\hat{y}_f(t-1), \dots, \hat{y}_f(t-n_d)]^T; \quad (40) \end{aligned}$$

$$\begin{aligned} \hat{y}_f(t) &= y(t) - [\hat{y}_f(t-1), \hat{y}_f(t-2), \dots, \\ &\hat{y}_f(t-n_d)]\hat{\mathbf{d}}(t); \quad (41) \end{aligned}$$

$$\begin{aligned} \hat{u}_f(t) &= u(t) - [\hat{u}_f(t-1), \hat{u}_f(t-2), \dots, \\ &\hat{u}_f(t-n_d)]\hat{\mathbf{d}}(t); \quad (42) \end{aligned}$$

$$\hat{\mathbf{b}}(t) = [\hat{b}_1(t), \hat{b}_2(t), \dots, \hat{b}_{n_b}(t)]^T; \quad (43)$$

$$\hat{\mathbf{d}}(t) = [\hat{d}_1(t), \hat{d}_2(t), \dots, \hat{d}_{n_d}(t)]^T; \quad (44)$$

$$\hat{\boldsymbol{\vartheta}}(t) = [\hat{\mathbf{b}}^T(t), \hat{\mathbf{d}}^T(t)]^T. \quad (45)$$

当新息长度  $p = 1$  时, F-MI-ESG 算法退化为 F-ESG 算法 (22) ~ (30). F-MI-ESG 算法 (35) ~ (45) 计算参数估计向量  $\hat{\boldsymbol{\vartheta}}(t)$  的步骤如下.

step 1: 令  $t = 1$ , 给定新息长度  $p$ . 置初值  $\hat{\boldsymbol{\vartheta}}(0) = \mathbf{1}_n/p_0, r(0) = 1, \hat{y}_f(t-i) = 1/p_0, \hat{u}_f(t-i) = 1/p_0, i = 1, 2, \dots, \max[n_b, n_d], p_0 = 10^6$ . 给定数据长度  $L_e$ .

step 2: 采集输入输出数据  $u(t)$  和  $y(t)$ , 按照式 (38) 组成构造堆积输出向量  $\mathbf{Y}(p, t)$ .

step 3: 利用式 (40) 构造滤波信息向量  $\hat{\boldsymbol{\varphi}}_f(t)$ , 由式 (39) 构造滤波信息矩阵  $\hat{\boldsymbol{\Phi}}_f(p, t)$ .

step 4: 根据式 (36) 计算新息向量  $\mathbf{E}(p, t)$ , 根据式 (37) 计算  $r(t)$ .

step 5: 利用式 (35) 更新参数估计向量  $\hat{\boldsymbol{\vartheta}}(t)$ .

step 6: 根据式 (45) 从  $\hat{\boldsymbol{\vartheta}}(t)$  中读取  $\hat{\mathbf{b}}(t)$  和  $\hat{\mathbf{d}}(t)$ .

step 7: 利用式 (41) 计算滤波输出  $\hat{y}_f(t)$ , 利用式 (42) 计算滤波输入  $\hat{u}_f(t)$ .

step 8: 如果  $t < L_e$ , 则  $t$  增加 1, 转至 step 2; 否则输出参数估计  $\hat{\boldsymbol{\vartheta}}(t)$ , 终止递推计算过程.

**引理 5** 对于 F-MI-ESG 算法 (35) ~ (45), 新息向量与残差为

$$\hat{\mathbf{V}}(p, t) := \mathbf{Y}(p, t) - \hat{\boldsymbol{\Phi}}_f^T(p, t)\hat{\boldsymbol{\vartheta}}(t),$$

满足

$$\hat{\mathbf{V}}(p, t) = \left[ \mathbf{I}_p - \frac{\hat{\boldsymbol{\Phi}}_f^T(p, t)\hat{\boldsymbol{\Phi}}_f(p, t)}{r(t)} \right] \mathbf{E}(p, t).$$

**证明** 利用式 (35), 有

$$\begin{aligned} \hat{\mathbf{V}}(p, t) &= \mathbf{Y}(p, t) - \hat{\boldsymbol{\Phi}}_f^T(p, t) \times \\ &\quad \left[ \hat{\boldsymbol{\vartheta}}(t-1) + \frac{\hat{\boldsymbol{\Phi}}_f(p, t)}{r(t)} \mathbf{E}(p, t) \right] = \\ &\quad \mathbf{Y}(p, t) - \hat{\boldsymbol{\Phi}}_f^T(p, t)\hat{\boldsymbol{\vartheta}}(t-1) - \\ &\quad \frac{\hat{\boldsymbol{\Phi}}_f^T(p, t)\hat{\boldsymbol{\Phi}}_f(p, t)}{r(t)} \mathbf{E}(p, t) = \\ &\quad \mathbf{E}(p, t) - \frac{\hat{\boldsymbol{\Phi}}_f^T(p, t)\hat{\boldsymbol{\Phi}}_f(p, t)}{r(t)} \mathbf{E}(p, t) = \\ &\quad \left[ \mathbf{I}_p - \frac{\hat{\boldsymbol{\Phi}}_f^T(p, t)\hat{\boldsymbol{\Phi}}_f(p, t)}{r(t)} \right] \mathbf{E}(p, t). \quad \square \end{aligned}$$

**引理 6** 对于基于滤波的多新息增广随机梯度算法 (35) ~ (45), 假设 SPE 条件成立, 定理 1 的条件成立, 那么下列不等式成立<sup>[15]</sup>:

$$1) \mathbf{I}_n - \frac{\hat{\boldsymbol{\Phi}}_f(p, t)\hat{\boldsymbol{\Phi}}_f^T(p, t)}{r(t)} \leq$$

$$\left[ 1 - \frac{p\alpha}{n\beta(t+N)+1} \right] \mathbf{I}_n, \text{ a.s.}$$

$$2) \|\hat{\boldsymbol{\Phi}}_f(p, t)\mathbf{Y}(p, t)\| \leq \left[ \mathbf{Y}^T(p, t)\hat{\boldsymbol{\Phi}}_f^T(p, t)\hat{\boldsymbol{\Phi}}_f(p, t)\mathbf{Y}(p, t) \right]^{1/2} \leq \sqrt{p\beta}\|\mathbf{Y}(p, t)\| \leq p\beta, \text{ a.s.}$$

**定理 2** 对于基于滤波的多新息增广随机梯度算法 (35) ~ (45), 假设 SPE 条件成立, 定理 1 的条件成立, 那么参数估计  $\hat{\boldsymbol{\vartheta}}(t)$  是有界收敛的.

对于 F-MI-ESG 算法的收敛性证明, 滤波为稳定是前提条件, 否则数据会发散. 在参数收敛的情况下, 使用估计的参数作为数字滤波器是可以保证稳定的, 每一时刻滤波器的参数也会进行更新, 因此只需要保证该算法收敛即可保证数字滤波器也是稳定的. 滤波器稳定, 持续激励条件映射数据有界, 参数估计收敛定理成立.

### 5 仿真分析

**例 1** ESG 算法和 F-ESG 算法.

考虑仿真对象 (FIR-MA 模型)

$$y(t) = B(z)u(t) + D(z)v(t),$$

$$B(z) = b_1z^{-1} + b_2z^{-2} =$$

$$-0.96z^{-1} - 0.37z^{-2},$$

$$D(z) = 1 + d_1z^{-1} + d_2z^{-2} =$$

$$1 - 0.20z^{-1} + 0.15z^{-2},$$

$$\boldsymbol{\vartheta} = [b_1, b_2, d_1, d_2]^T =$$

$$[-0.96, -0.37, -0.20, 0.15]^T.$$

仿真时, 输入  $u(t)$  采用零均值单位方差不相关的可测随机信号序列,  $v(t)$  采用零均值方差为  $\sigma^2$  的白噪声序列. 考虑 3 种不同噪声水平, 噪声方差分别为  $\sigma^2 = 0.50^2, \sigma^2 = 1.50^2$  和  $\sigma^2 = 2.00^2$ . 利用输入输出数据  $u(t)$  和  $y(t)$ , 应用增广随机梯度 (ESG) 算法 (10) ~ (16) 和基于滤波的增广随机梯度 (F-ESG) 算法 (22) ~ (30) 估计系统的参数, 不同噪声方差下参数估计及其误差如表 1 和表 2 所示. 参数估计误差  $\delta(t) := \|\hat{\boldsymbol{\vartheta}}(t) - \boldsymbol{\vartheta}\|/\|\boldsymbol{\vartheta}\|$  随  $t$  的变化曲线如图 1 和图 2 所示.

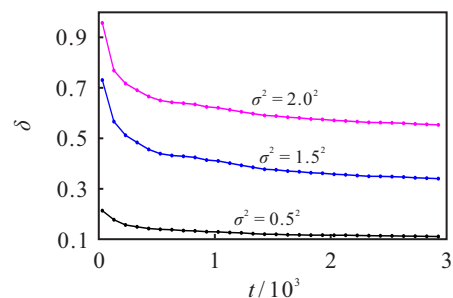


图 1 ESG 算法参数估计误差  $\delta(t)$  随  $t$  变化曲线

表1 ESG算法参数估计及其误差( $\sigma^2 = 0.50^2, \sigma^2 = 1.50^2, \sigma^2 = 2.00^2$ )

$\sigma^2$	$t$	$b_1$	$b_2$	$d_1$	$d_2$	$\delta/\%$
0.50 <sup>2</sup>	100	-1.010 66	-0.245 24	-0.077 54	0.218 75	18.376 71
	200	-0.991 30	-0.265 44	-0.078 64	0.214 11	16.562 69
	500	-0.967 43	-0.300 21	-0.079 71	0.206 35	14.189 37
	1 000	-0.970 20	-0.316 29	-0.087 51	0.208 56	13.043 49
	2 000	-0.965 08	-0.336 57	-0.093 11	0.201 74	11.662 15
	3 000	-0.962 88	-0.341 03	-0.096 89	0.198 74	11.117 26
1.50 <sup>2</sup>	100	-1.308 71	0.042 49	-0.431 55	0.408 81	60.649 69
	200	-1.266 81	0.027 67	-0.347 29	0.342 70	52.680 57
	500	-1.215 54	-0.011 54	-0.288 93	0.295 06	44.575 58
	1 000	-1.197 84	-0.038 55	-0.272 26	0.288 08	41.246 15
	2 000	-1.169 63	-0.073 15	-0.259 09	0.250 60	36.048 90
	3 000	-1.155 86	-0.086 57	-0.253 34	0.237 00	33.936 74
2.00 <sup>2</sup>	100	-1.436 47	0.217 57	-0.470 53	0.447 73	80.922 69
	200	-1.401 53	0.204 00	-0.369 10	0.363 68	73.079 42
	500	-1.356 82	0.169 44	-0.301 27	0.304 84	65.619 54
	1 000	-1.338 74	0.144 12	-0.281 14	0.294 39	62.307 53
	2 000	-1.310 68	0.111 47	-0.266 00	0.249 09	57.370 92
	3 000	-1.296 40	0.097 73	-0.259 67	0.233 22	55.268 46
真值		-0.960 00	-0.370 00	-0.200 00	0.150 00	

表2 F-ESG算法参数估计及其误差( $\sigma^2 = 0.50^2, \sigma^2 = 1.50^2, \sigma^2 = 2.00^2$ )

$\sigma^2$	$t$	$b_1$	$b_2$	$d_1$	$d_2$	$\delta/\%$
0.50 <sup>2</sup>	100	-0.545 13	-0.441 59	-0.057 29	-0.000 78	44.335 36
	200	-0.583 46	-0.459 41	-0.038 03	-0.003 57	42.195 79
	500	-0.648 13	-0.475 42	-0.021 96	-0.023 93	38.979 57
	1 000	-0.687 75	-0.482 67	-0.015 47	-0.027 28	36.858 77
	2 000	-0.723 43	-0.490 20	-0.007 70	-0.036 60	35.617 76
	3 000	-0.741 80	-0.497 53	0.000 05	-0.040 42	35.359 08
1.50 <sup>2</sup>	100	-0.943 19	-0.320 19	-0.517 40	0.229 50	31.300 15
	200	-0.945 80	-0.324 16	-0.428 63	0.217 61	22.969 84
	500	-0.947 41	-0.356 08	-0.307 17	0.208 04	11.646 29
	1 000	-0.947 55	-0.359 57	-0.267 76	0.192 13	7.690 55
	2 000	-0.953 39	-0.377 02	-0.221 54	0.157 90	2.350 24
	3 000	-0.954 01	-0.385 84	-0.195 96	0.147 21	1.665 49
2.00 <sup>2</sup>	100	-0.916 32	-0.345 42	-0.546 75	0.322 11	36.867 42
	200	-0.918 78	-0.346 10	-0.440 22	0.286 60	26.485 61
	500	-0.920 44	-0.372 58	-0.306 57	0.260 17	14.953 73
	1 000	-0.920 82	-0.372 50	-0.263 02	0.234 30	10.610 01
	2 000	-0.926 88	-0.385 79	-0.214 06	0.187 51	5.130 85
	3 000	-0.927 60	-0.392 40	-0.187 25	0.170 96	4.383 36
真值		-0.960 00	-0.370 00	-0.200 00	0.150 00	

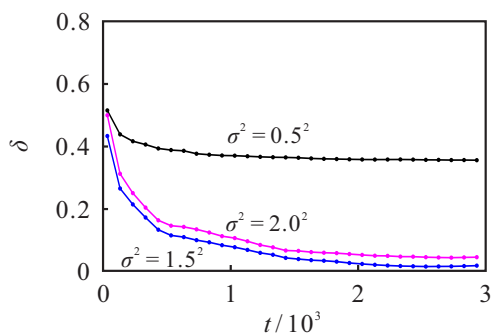


图2 F-ESG算法参数估计误差 $\delta(t)$ 随 $t$ 变化曲线

由表1、表2和图1、图2可知:

1) 随着 $t$ 的增加,ESG算法和F-ESG算法给出的参数估计误差总体趋势是减小的.

2) 在相同条件下,F-ESG算法比ESG算法有更小的参数估计误差.即与ESG算法相比,F-ESG算法具有更高的精度.

例2 F-ESG算法和F-MI-ESG算法.

考虑仿真对象

$$y(t) = B(z)u(t) + D(z)v(t),$$

$$B(z) = b_1z^{-1} + b_2z^{-2} = 0.67z^{-1} - 0.42z^{-2},$$

$$D(z) = 1 + d_1z^{-1} + d_2z^{-2} =$$

$$1 - 0.21z^{-1} + 0.15z^{-2},$$

$$\vartheta = [b_1, b_2, d_1, d_2]^T =$$

$$[0.67, -0.42, -0.21, 0.15]^T.$$

仿真时,输入 $u(t)$ 采用零均值单位方差不相关

可测随机信号序列,  $v(t)$  采用零均值方差为  $\sigma^2 = 0.50^2$  白噪声序列, 应用基于滤波的增广随机梯度算法 (22)~(30) (即在 F-MI-ESG 算法中  $p = 1$ ) 和基于滤波的多新息增广随机梯度算法 (35)~(45) 估计系统的参数, F-ESG 参数估计和不同新息长度下的 F-

MI-ESG 参数估计及误差如表 3 所示, 参数估计误差  $\delta(t) := \|\hat{\vartheta}(t) - \vartheta\|/\|\vartheta\|$  随  $t$  的变化曲线如图 3 所示 ( $\sigma^2 = 0.50^2$ ). F-MI-ESG 算法在方差  $\sigma^2 = 0.50^2$ 、新息长度  $p = 6$  时的参数估计值随  $t$  变化曲线如图 4 所示 ( $\sigma^2 = 0.50^2, p = 6$ ).

表 3 F-MI-ESG 算法参数估计及其误差 ( $\sigma^2 = 0.50^2$ )

$p$	$t$	$b_1$	$b_2$	$d_1$	$d_2$	$\delta/\%$
1	100	0.401 91	-0.194 81	-0.380 48	0.062 23	47.990 26
	200	0.438 25	-0.206 08	-0.366 16	0.067 58	43.454 09
	500	0.488 37	-0.252 04	-0.368 59	0.063 03	36.842 47
	1 000	0.516 10	-0.262 59	-0.363 86	0.071 01	33.656 34
	2 000	0.534 98	-0.285 56	-0.368 68	0.073 25	31.205 38
	3 000	0.545 82	-0.297 05	-0.368 51	0.075 00	29.761 94
2	100	0.548 73	-0.249 38	-0.381 11	0.112 64	32.811 47
	200	0.575 08	-0.268 88	-0.357 57	0.118 05	28.102 94
	500	0.611 12	-0.340 10	-0.355 31	0.100 57	21.973 77
	1 000	0.628 89	-0.343 22	-0.341 27	0.116 33	19.367 45
	2 000	0.634 90	-0.369 17	-0.345 41	0.121 30	18.222 23
	3 000	0.640 87	-0.380 29	-0.342 03	0.124 48	17.216 56
3	100	0.602 27	-0.271 00	-0.361 08	0.140 16	26.804 22
	200	0.619 06	-0.295 40	-0.331 85	0.143 52	21.842 24
	500	0.646 94	-0.381 84	-0.324 65	0.113 26	15.434 27
	1 000	0.658 25	-0.372 59	-0.303 38	0.135 74	12.784 80
	2 000	0.658 15	-0.397 81	-0.308 91	0.142 57	12.302 27
	3 000	0.662 81	-0.407 00	-0.304 21	0.146 29	11.474 58
4	100	0.645 16	-0.296 58	-0.289 99	0.164 22	18.013 81
	200	0.645 82	-0.321 01	-0.252 19	0.161 63	13.332 33
	500	0.675 15	-0.429 72	-0.233 81	0.096 78	7.132 94
	1 000	0.669 28	-0.387 00	-0.207 24	0.138 62	4.210 72
	2 000	0.668 32	-0.417 46	-0.230 80	0.150 08	2.526 79
	3 000	0.674 84	-0.424 67	-0.229 50	0.154 92	2.549 50
真值		0.670 00	-0.420 00	-0.210 00	0.150 00	

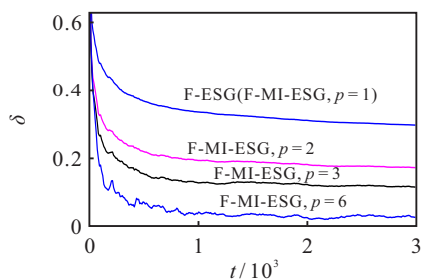


图 3 F-MI-ESG 算法参数估计误差  $\delta(t)$  随  $t$  变化曲线

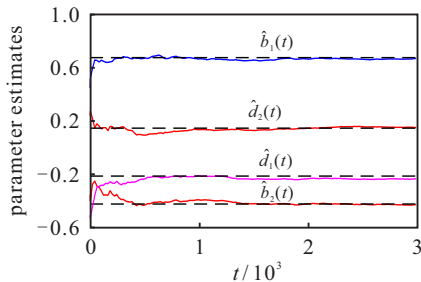


图 4 F-MI-ESG 算法参数估计值随  $t$  变化曲线

从表 3、图 3 和图 4 可以看出: F-MI-ESG 算法的参数估计误差随  $t$  的增加而减小, 且 F-MI-ESG 估计比 F-ESG 估计具有更高的精度. 在相同噪声方差下,

随着新息长度  $p$  增加, F-MI-ESG 算法的参数估计误差越来越小. 此外, 当新息长度  $p = 6$  时, 估计值非常接近真实值. 因此, 适当的新息长度  $p$  使得 F-MI-ESG 算法具有良好的估计精度.

对于模型验证, 选择  $p = 6, t = 3000$  时, 表 3 中倒数第 2 行的参数估计作为最终的估计模型, 即

$$\begin{aligned} \hat{y}(t) &= y(t) - \hat{y}_f(t) + \hat{B}(z)\hat{u}_f(t) = \\ y(t) - \hat{y}_f(t) + 0.67484\hat{u}_f(t-1) - 0.42467\hat{u}_f(t-2), \\ \hat{y}_f(t) &= y(t) - [\hat{y}_f(t-1) + \hat{y}_f(t-2)]\hat{d}(t) = \\ y(t) - 0.22950\hat{y}_f(t-1) + 0.15492\hat{y}_f(t-2), \\ \hat{u}_f(t) &= u(t) - [\hat{u}_f(t-1) + \hat{u}_f(t-2)]\hat{d}(t) = \\ u(t) - 0.22950\hat{u}_f(t-1) + 0.15492\hat{u}_f(t-2). \end{aligned}$$

将  $t = L_e + 1 = 3001$  到  $t = L_e + L_r = 4000$  的 1000 组输入输出数据  $u(t)$  和  $y(t)$  代入上式, 可以递推计算出估计模型的输出  $\hat{y}(t)$ , 进而计算估计模型输出误差均方根为

$$\text{Error}(L_r) = \left[ \frac{1}{L_r} \sum_{t=L_e+1}^{L_e+L_r} [\hat{y}(t) - y(t)]^2 \right]^{\frac{1}{2}} = 0.54294.$$

系统输出  $y(t)$  和模型输出  $\hat{y}(t)$  及其误差  $\hat{y}(t) - y(t)$  随  $t$  的变化情况如图5所示 ( $\sigma^2 = 0.50^2, p = 6$ ).

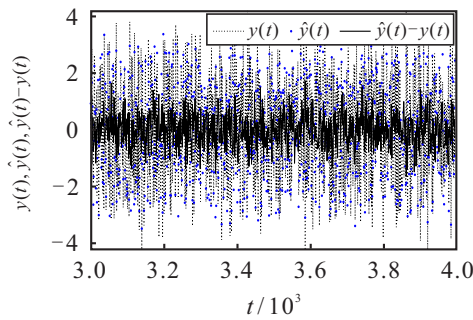


图5 F-MI-ESG算法获得的预测模型输出  $\hat{y}(t)$  与系统输出  $y(t)$  及其误差随  $t$  变化曲线

从图5可以看出: 模型输出很接近系统输出, 模型输出误差很小, 接近噪声的标准差  $\sigma$ , 表明估计模型能很好地捕捉系统动态.

## 6 结论

本文使用数据滤波技术提出了基于滤波的随机梯度(F-ESG)算法, 并在此基础上推导了基于滤波的多新息随机梯度算法(F-MI-ESG算法). 仿真结果表明, F-ESG算法比随机梯度(ESG)算法有更高的参数估计精度, 而F-MI-ESG算法可以加快收敛速度, 提高参数估计的精度. 所提出方法的基本思想可以扩展到研究其他线性和非线性系统, 并应用于其他领域.

## 参考文献(References)

- [1] 初蕾, 朱善良, 王明新, 等. 具有有限时间输出约束的切换非线性时滞系统的多维泰勒网自适应控制[J]. 控制与决策, 2022, 37(2): 361-369.  
(Chu L, Zhu S L, Wang M X, et al. Multi-dimensional Taylor network adaptive control for switched time-delay systems with finite-time output constraints[J]. Control and Decision, 2022, 37(2): 361-369.)
- [2] 蔡赛男, 宋卫星, 班利明, 等. 基于鲸鱼算法优化LSSVM的滚动轴承故障诊断[J]. 控制与决策, 2022, 37(1): 230-236.  
(Cai S N, Song W X, Ban L M, et al. Fault diagnosis method of rolling bearing based on LSSVM optimized by whale optimization algorithm[J]. Control and Decision, 2022, 37(1): 230-236.)
- [3] 王印松, 孙天舒. 一种基于证据融合的执行器故障诊断方法[J]. 控制与决策, 2022, 37(8): 2026-2032.  
(Wang Y S, Sun T S. A method of actuator fault diagnosis based on evidence fusion[J]. Control and Decision, 2022, 37(8): 2026-2032.)
- [4] 张霄, 丁锋. 双线性状态空间系统的状态观测器设计[J]. 控制与决策, 2023, 38(1): 274-280.  
(Zhang X, Ding F. State observers for bilinear state-space systems[J]. Control and Decision, 2023, 38(1): 274-280.)
- [5] Wahlberg B, Ljung L. Algorithms and performance analysis for stochastic Wiener system identification[J]. IEEE Control Systems Letters, 2018, 2(3): 471-476.
- [6] Wang T, Hu M, Zhao Y L. Adaptive tracking control of FIR systems under binary-valued observations and recursive projection identification[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2021, 51(9): 5289-5299.
- [7] Munker T, Nelles O. Nonlinear system identification with regularized local FIR model networks[J]. Engineering Applications of Artificial Intelligence, 2018, 67: 345-354.
- [8] Tao T Y, Wang B, Wang X H. Parameter and time delay estimation algorithm based on gradient pursuit for multi-input C-ARMA systems[J]. Control and Decision, 2022, 37(8): 2085-2090.
- [9] Fan S J, Xu L, Ding F, et al. Correlation analysis-based stochastic gradient and least squares identification methods for errors-in-variables systems using the multiinnovation[J]. International Journal of Control, Automation and Systems, 2021, 19(1): 289-300.
- [10] Fan Y M, Liu X M. Data filtering-based multi-innovation forgetting gradient algorithms for input nonlinear FIR-MA systems with piecewise-linear characteristics[J]. Journal of the Franklin Institute, 2021, 358(18): 9818-9840.
- [11] Ma P, Wang L. Filtering-based recursive least squares estimation approaches for multivariate equation-error systems by using the multiinnovation theory[J]. International Journal of Adaptive Control and Signal Processing, 2021, 35(9): 1898-1915.
- [12] 丁锋. 系统辨识: 多新息辨识理论与方法[M]. 北京: 科学出版社, 2016.  
(Ding F. System identification: Theory and method of multi-innovation identification[M]. Beijing: Science Press, 2016.)
- [13] Liu S Y, Zhang Y L, Xu L, et al. Extended gradient-based iterative algorithm for bilinear state-space systems with moving average noises by using the filtering technique[J]. International Journal of Control, Automation and Systems, 2021, 19(4): 1597-1606.
- [14] Wang D Q. Least Squares-based recursive and iterative estimation for output error moving average systems using data filtering[J]. IET Control Theory & Applications, 2011, 5(14): 1648-1657.
- [15] 丁锋. 系统辨识-辨识方法性能分析[M]. 北京: 科学出版社, 2014.  
(Ding F. Performance analysis of system identification-identification method[M]. Beijing: Science Press, 2014.)

## 作者简介

丁锋(1963—), 男, 教授, 博士生导师, 从事系统辨识、自适应控制等研究, E-mail: fding@jiangnan.edu.cn;

郑嘉芸(1999—), 女, 硕士生, 从事系统辨识的研究, E-mail: jyzheng12@126.com;

张霄(1994—), 女, 副教授, 博士, 从事系统辨识、状态估计等研究, E-mail: xzhang@jiangnan.edu.cn;

徐玲(1976—), 女, 教授, 博士, 从事系统建模、系统辨识等研究, E-mail: lingxu0848@163.com.