Decomposition-based recursive least squares identification methods for multivariate pseudo-linear systems using the multi-innovation

Ping Ma\textsuperscript{a}, Feng Ding\textsuperscript{a,b} and Quanmin Zhu\textsuperscript{c}

\textsuperscript{a}Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education), School of Internet of Things Engineering, Jiangnan University, Wuxi 214122, PR China; \textsuperscript{b}College of Automation and Electronic Engineering, Qingdao University of Science and Technology, Qingdao 266042, PR China; \textsuperscript{c}Department of Engineering Design and Mathematics, University of the West of England, Bristol, England

ARTICLE HISTORY
Received 3 July 2017
Accepted 21 January 2018

KEYWORDS
Parameter estimation; least squares; multi-innovation; decomposition technique; multivariate system

ABSTRACT
This paper studies the parameter estimation algorithms of multivariate pseudo-linear autoregressive systems. A decomposition-based recursive generalised least squares algorithm is deduced for estimating the system parameters by decomposing the multivariate pseudo-linear autoregressive system into two subsystems. In order to further improve the parameter accuracy, a decomposition based multi-innovation recursive generalised least squares algorithm is developed by means of the multi-innovation theory. The simulation results confirm that these two algorithms are effective.

1. Introduction
System identification is the modelling theory and method for researching on static systems and dynamic systems (Anfinsen, Diagne, Aamo, & Krstic, 2017; Bravo, Suarez, Vasallo, & Alamo, 2016; Zhu, Yu, & Zhao, 2017). Mathematical model can be established by the mechanism analysis method and the statistical method (Bessaoudi, Ben Hmida, & Hsieh, 2017; Escobar & Poznyak, 2015; Zhu, Wang, Zhao, Li, & Billings, 2015). The mechanism analysis method uses equations of the law of the system motion to obtain the system model structure. Model parameters can be gained by measuring, as well as some kinds of identification methods by collecting the system input and output data (Ding, Xu, & Zhu, 2016; Ma, Ding, Alsaeedi, & Hayat, 2018; Xu & Ding, 2017). In recent years, some system identification methods have been presented by researchers all over the world. Isaksson, Sjoberg, Tornqvist, Ljung, and Kok (2015) proposed an established method for grey-box identification by using the maximum-likelihood estimation and the extended Kalman filtering. Wang and Ding (2016) presented a recursive generalised extended least squares algorithm and a generalised extended stochastic gradient algorithm for identifying the parameters of a class of nonlinear systems. Ma and Ding (2017) derived a filtering-based generalised stochastic gradient algorithm for multivariate pseudo-linear autoregressive systems by means of the data filtering technique.

Some typical industrial processes are multivariate systems in essence (Milad & Javad, 2017; Ramezani, Arefi, Zargarzadeh, & Jahed-Motlagh, 2017; Tohru & Hajime, 2016). Multivariate systems have the characteristics of high dimensions, many input and output variables and lots of system parameters, and these factors result in heavy computational burden of identification algorithms (Mahnaz, 2017; Sergey & Svyatoslav, 2016; Theodoreakopoulos & Rovithakis, 2016). The decomposition technique is an effective method to separate a large-scale system into several small-sized subsystems (Ding, Wang, Xu, Tasawar, & Ahmed, 2017; Ding, Wang, Xu, & Wu, 2017; Wang, Mao, & Ding, 2017). Applying the decomposition technique to the multivariate system can reduce the computational burden. Ma, Xiong, Chen, and Ding (2017) developed a modified Kalman filter based hierarchical least squares algorithm for multi-input multi-output Hammerstein systems. The highlight of that paper is to adopt the hierarchical identification to decompose the system into two fictitious subsystems, one containing the unknown parameters in the non-linear block and the other containing the unknown parameters in the linear subsystems.

The multi-innovation identification theory can improve the parameter estimation accuracy by expanding the identification innovation from a scalar innovation to an innovation vector, or from a vector innovation to an innovation matrix (Hu, Liu, & Zhou, 2014; Shi, Wang, & Ji, 2016b; Li & Liu, 2018). In this aspect, Meng (2017)
derived a multi-innovation stochastic gradient algorithm for identifying the parameters of bilinear systems based on the least squares principle and the multi-innovation identification theory. Ding, Wang, Mao, and Xu (2017) developed a filtering-based multi-innovation gradient algorithm for a linear state space system with time-delay, and analysed that the parameter estimates given by the presented algorithms converge to their true values under the persistent excitation conditions. Chen, Ding, Alsaeedi, and Hayat (2017) derived a filtering-based maximum likelihood multi-innovation extended gradient algorithm to estimate the parameters of controlled autoregressive moving average systems by replacing the unmeasurable variables in the information vectors with zero mean, $C(z) \in \mathbb{R}^{m \times m}$ is a white noise vector with zero mean, $C(z) \in \mathbb{R}^{m \times m}$ is a polynomial matrix in the unit backward shift operator $[z^{-1}y(t) = y(t - 1)]$, where $y(t) := \{y_1(t), y_2(t), \ldots, y_m(t)\}^T \in \mathbb{R}^m$ is the system output vector, $\Phi_s(t) \in \mathbb{R}^{m \times n}$ is the system information matrix consisting of the input–output data, $\theta \in \mathbb{R}^n$ is the system parameter vector to be identified, $v(t) := [v_1(t), v_2(t), \ldots, v_m(t)]^T \in \mathbb{R}^m$ is a white noise vector with zero mean.

Consider the following multivariate pseudo-linear autoregressive system:

$$y(t) = \Phi_s(t)\theta + C^{-1}(z)v(t),$$

where $y(t) := \{y_1(t), y_2(t), \ldots, y_m(t)\}^T \in \mathbb{R}^m$ is the system output vector, $\Phi_s(t) \in \mathbb{R}^{m \times n}$ is the system information matrix consisting of the input–output data, $\theta \in \mathbb{R}^n$ is the system parameter vector to be identified, $v(t) := [v_1(t), v_2(t), \ldots, v_m(t)]^T \in \mathbb{R}^m$ is a white noise vector with zero mean, $C(z) \in \mathbb{R}^{m \times m}$ is a polynomial matrix in the unit backward shift operator $[z^{-1}y(t) = y(t - 1)]$.

Without loss of generality, assume that the orders $m$, $n$ and $n_c$ are known, and $y(t) = 0$, $\Phi_s(t) = 0$ and $v(t) = 0$ for $t \leq 0$.

Define the related noise vector $w(t)$, the noise model parameter matrix $\theta_n$ and the noise model information vector $\psi(t)$ as

$$w(t) := C^{-1}(z)v(t) \in \mathbb{R}^m,$$

$$\theta_n := [C_1, C_2, \ldots, C_n]^T \in \mathbb{R}^{(mn_c) \times m},$$

$$\psi(t) := [-w^T(t - 1), -w^T(t - 2), \ldots, -w^T(t - n_c)]^T \in \mathbb{R}^{mn_c}.$$ (2)

From Equation (2), we have

$$w(t) = [I_m - C(z)]w(t) + v(t),$$

$$= (-C_1z^{-1} - C_2z^{-2} - \cdots - C_nz^{-n_c})w(t) + v(t)$$

$$= -C_1w(t - 1) - C_2w(t - 2) - \cdots - C_nw(t - n_c) + v(t)$$

$$= \theta_n^T\psi(t) + v(t).$$ (3)

Equation (3) is the noise model. Therefore, Equation (1) can be rewritten as

$$y(t) = \Phi_s(t)\theta + w(t)$$

$$= \Phi_s(t)\theta + \theta_n^T\psi(t) + v(t).$$ (4)

This identification model contains a system model parameter vector $\theta$ and a noise model parameter matrix $\theta_n$. $\Phi_s(t)$ is the system information matrix consisting of the input–output data, and thus is known, $\psi(t)$ is the noise model information vector consisting of the unmeasurable noise items $w(t - i)$, $i = 1, 2, \ldots, n_c$, so it is unknown.

The rest of this paper is organised as follows. Section 2 describes a multivariate pseudo-linear autoregressive system and decomposes it into two sub-identification models. Section 3 derives a M-D-RGLS algorithm. Section 4 obtains a M-D-MI-RGLS algorithm. Section 5 provides an example for illustrating the results in this paper. Finally, we offer some concluding remarks in Section 6.

2. System description and identification model

First of all, we give some notations in this paper. $I_m$ denotes an identity matrix of size $m \times m$; $I_n$ stands for an $n$-dimensional column vector whose elements are 1, that is $I_n := [1, 1, \ldots, 1]^T \in \mathbb{R}^n$; $I_m \times n$ represents a matrix of size $m \times n$ whose elements are 1; the norm of a matrix $X$ is defined by $\|X\|^2 := \text{tr}[XX^T]$, the superscript $T$ stands for the matrix/vector transpose.
3. The M-D-RGLS algorithm

In this section, we derive a new decomposition-based recursive least squares algorithm. Based on the identification models in Equations (7) and (8), define two quadratic functions:

\[
J_1(\theta) := \sum_{j=1}^{t} \| y_1(j) - \Phi_s(j)\theta \|^2,
\]

\[
J_2(\hat{\theta}_n) := \sum_{j=1}^{t} \| w(j) - \hat{\theta}_n^T\psi(j) \|^2.
\]

Using the least squares principle (Shi, Wang, & Ji, 2016a; Wang & Zhang, 2015; Xu & Ding, 2017; Zhang, Ding, Alsaeedi, & Hayat, 2017) and minimising \( J_1(\theta) \) and \( J_2(\hat{\theta}_n) \), we have two recursive relationships:

\[
\hat{\theta}(t) = \hat{\theta}(t-1) + P_1(t)\Phi_s^T(t)[y_1(t) - \Phi_s(t)\hat{\theta}(t-1)],
\]

\[
P_1^{-1}(t) = P_1^{-1}(t-1) + \Phi_s^T(t)\Phi_s(t),
\]

\[
\hat{\theta}_n(t) = \hat{\theta}_n(t-1) + P_2(t)\psi(t)[w(t) - \hat{\theta}_n^T(t-1)\psi(t)]^T,
\]

\[
P_2^{-1}(t) = P_2^{-1}(t-1) + \psi(t)\psi^T(t).
\]

When calculating the parameter estimations \( \hat{\theta}(t) \) and \( \hat{\theta}_n(t) \), in order to avoid the inverse of error covariance matrices \( P_1(t) \in \mathbb{R}^{n \times n} \) and \( P_2(t) \in \mathbb{R}^{m \times m} \), we use the matrix inversion formula

\[
(A + BC)^{-1} = A^{-1} - A^{-1}B(I + CA^{-1}B)^{-1}CA^{-1},
\]

and introduce a gain matrix \( L_1(t) := P_1(t)\Phi_s^T(t) \in \mathbb{R}^{n \times m} \) and a gain vector \( L_2(t) := P_2(t)\psi(t) \in \mathbb{R}^m \). The algorithm in Equations (9)–(12) can be converted to an equivalent form:

\[
\hat{\theta}(t) = \hat{\theta}(t-1) + L_1(t)[y_1(t) - \Phi_s(t)\hat{\theta}(t-1)],
\]

\[
L_1(t) = P_1(t-1)\Phi_s^T(t)[I_m + \Phi_s(t)P_1(t-1)\Phi_s^T(t)]^{-1},
\]

\[
P_1(t) = [I_n - L_1(t)\Phi_s(t)]P_1(t-1),
\]

\[
\hat{\theta}_n(t) = \hat{\theta}_n(t-1) + L_2(t)[w(t) - \hat{\theta}_n^T(t-1)\psi(t)]^T,
\]

\[
L_2(t) = P_2(t-1)\psi(t)[1 + \psi^T(t)P_2(t-1)\psi(t)]^{-1}.
\]
\( P_2(t) = [I_{mn} - L_2(t) \hat{\psi}^T(t)] P_2(t - 1) \) \hspace{1cm} (18)

Here, the difficulty of identification is that \( \hat{\psi}(t) \) contains the unmeasurable noise terms \( w(t - i) \), the estimates \( \hat{\theta}(t) \) and \( \hat{\theta}_n(t) \) in Equations (13)–(18) are impossible to compute. The solution is to replace the unmeasurable variables \( w(t - i) \) with their corresponding estimates \( \hat{w}(t - i) \). Define the estimate of \( \hat{\psi}(t) \) as

\[
\hat{\psi}(t) := [-\hat{w}^T(t - 1), -\hat{w}^T(t - 2), \ldots, -\hat{w}^T(t - n_c)]^T.
\]

(19)

From Equations (5) and (6), we can compute the estimates \( \hat{y}_1(t) \) and \( \hat{w}(t) \) by replacing the unknown parameters \( \theta_n \) and \( \theta \) with the estimates \( \hat{\theta}_n(t - 1) \) and \( \hat{\theta}(t - 1) \):

\[
\hat{y}_1(t) = y(t) - \hat{\theta}_n(t - 1) \hat{\psi}(t), \quad (20)
\]

\[
\hat{w}(t) = y(t) - \Phi_s(t) \hat{\theta}(t - 1). \quad (21)
\]

Furthermore, replacing \( y_1(t) \), \( w(t) \) and \( \psi(t) \) in Equations (13)–(18) with their corresponding estimates \( \hat{y}_1(t) \), \( \hat{w}(t) \) and \( \hat{\psi}(t) \), and combining the Equations (19)–(21), we can obtain the M-D-RGLS algorithm for estimating the parameter vector \( \hat{\theta} \):

\[
\hat{\theta}(t) = \hat{\theta}(t - 1) + L_1(t)[\hat{y}_1(t) - \Phi_s(t) \hat{\theta}(t - 1)], \quad (22)
\]

\[
P_1(t) = [I_n - L_1(t) \Phi_s(t)] P_1(t - 1), \quad (24)
\]

\[
\hat{\theta}_n(t) = \hat{\theta}_n(t - 1) + L_2(t)[\hat{w}(t) - \hat{\theta}_n(t - 1) \hat{\psi}(t)]^T, \quad (25)
\]

\[
P_2(t) = [I_{mn} - L_2(t) \hat{\psi}^T(t)] P_2(t - 1), \quad (27)
\]

\[
\hat{y}_1(t) = y(t) - \hat{\theta}_n(t - 1) \hat{\psi}(t), \quad (28)
\]

\[
\hat{w}(t) = y(t) - \Phi_s(t) \hat{\theta}(t - 1), \quad (29)
\]

\[
\hat{\psi}(t) = [-\hat{w}^T(t - 1), -\hat{w}^T(t - 2), \ldots, -\hat{w}^T(t - n_c)]^T. \quad (30)
\]

The procedures of computing the parameter estimation vector \( \hat{\theta}(t) \) by the M-D-RGLS algorithm in Equations (22)–(30) are listed as follows:

1. Let \( t = 1 \), set the initial values \( \hat{\theta}(0) = 1_n/p_0, \hat{\theta}_n(0) = 1_{(mn)_x_m}/p_0, \quad P_1(0) = p_0 I_n, \quad P_2(0) = p_0 I_{mn} \).

2. Collect the observation data \( \Phi_s(t) \) and \( y(t) \).

3. Compute gain matrices \( L_1(t) \) and \( L_2(t) \) using Equations (23) and (26).

4. Compute covariance matrices \( P_1(t) \) and \( P_2(t) \) using Equations (24) and (27).

5. Compute \( \hat{y}_1(t) \) and \( \hat{w}(t) \) using Equations (28) and (29).

6. Update the parameter estimates \( \hat{\theta}(t) \) and \( \hat{\theta}_n(t) \) using Equations (22) and (25).

7. Increase \( t \) by 1, go to Step 2.

The flowchart of computing \( \hat{\theta}(t) \) in the M-D-RGLS algorithm is shown in Figure 2.

4. The M-D-MI-RGLS algorithm

The multi-innovation identification theory can extract more useful information from observation data to improve the parameter estimation accuracy (Chen, Ding, Xu, Hayat, & Alsaedi, 2017; Mao & Ding, 2016; Pan, Jiang, Wan, & Ding, 2017). Based on the M-D-RGLS algorithm in Equations (22)–(30), according to the multi-innovation identification theory, we introduce an innovation length \( p \) to expand the innovation vector to a large innovation vector/matrix. Define the stacked information matrices \( \Gamma_1(p, t) \) and \( \Gamma_2(p, t) \), the stacked output vector \( Y(p, t) \) and the stacked noise output matrix \( W(p, t) \) as

\[
\Gamma_1(p, t) := [\Phi_s^T(t), \Phi_s^T(t - 1), \ldots, \Phi_s^T(t - p + 1)] \in \mathbb{R}^{n \times (mp)};
\]

\[
\Gamma_2(p, t) := [\hat{\psi}(t), \hat{\psi}(t - 1), \ldots, \hat{\psi}(t - p + 1)] \in \mathbb{R}^{(mn)_x_p},
\]
Normally, we reach an agreement that the estimates $\hat{\theta}_n(t)$ and $\hat{\theta}_n(t-1)$ at time $(t-1)$ are closer to the true values $\theta$ and $\theta_n$ than the estimates $\hat{\theta}(t-i)$ and $\hat{\theta}_n(t-i)$ at time $(t-i) (i \geq 2)$. Therefore, replacing the terms $\hat{\theta}(t-i)$ and $\hat{\theta}_n(t-i) (i \geq 2)$ with $\hat{\theta}(t-1)$ and $\hat{\theta}_n(t-1)$ in Equations (31) and (32), the large innovation vector $E_1(p, t)$ and the innovation matrix $E_2(p, t)$ can be modified into

$$E_1(p, t) := \begin{bmatrix}
\hat{y}_1(t) - \Phi_1(t)\hat{\theta}(t-1) \\
\hat{y}_1(t-1) - \Phi_1(t-1)\hat{\theta}(t-2) \\
\vdots \\
\hat{y}_1(t-p+1) - \Phi_1(t-p+1)\hat{\theta}(t-p)
\end{bmatrix} \in \mathbb{R}^{mp},$$

and

$$E_2(p, t) := \begin{bmatrix}
\hat{y}_1(t) - \Phi_1(t)\hat{\theta}(t-1) \\
\hat{y}_1(t-1) - \Phi_1(t-1)\hat{\theta}(t-2) \\
\vdots \\
\hat{y}_1(t-p+1) - \Phi_1(t-p+1)\hat{\theta}(t-p)
\end{bmatrix} \in \mathbb{R}^{mp}.\quad (31)$$

In addition, define new gain matrices $L_3(t) := P_1(t)\Gamma_1(p, t) \in \mathbb{R}^{m \times (mp)}$ and $L_4(t) := P_2(t)\Gamma_2(p, t) \in \mathbb{R}^{(mn) \times p}$. We can obtain the following M-D-MI-RGLS algorithm with an innovation length $p$:

$$\hat{\theta}(t) = \hat{\theta}(t-1) + L_3(t)E_1(p, t),$$

$$E_1(p, t) = Y(p, t) - \Gamma_1^T(p, t)\hat{\theta}(t-1),$$

$$L_3(t) = P_1(t-1)\Gamma_1(p, t)[I_{mp} + \Gamma_1^T(p, t)P_1(t-1)\Gamma_1(p, t)]^{-1},$$

$$P_1(t) = [I_n - L_3(t)\Gamma_1^T(p, t)]P_1(t-1),$$

$$Y(p, t) = [\tilde{y}_1(t), \tilde{y}_1^T(t-1), \ldots, \tilde{y}_1^T(t-p+1)]^T,$$

$$\Gamma_1(p, t) = [\Phi_1^T(t), \Phi_1^T(t-1), \ldots, \Phi_1^T(t-p+1)].$$

$$\hat{\theta}_n(t) = \hat{\theta}_n(t-1) + L_4(t)E_2(p, t),$$

$$E_2(p, t) = W(p, t) - \Gamma_2^T(p, t)\hat{\theta}_n(t-1),$$

$$L_4(t) = P_2(t-1)\Gamma_2(p, t)[I_p + \Gamma_2^T(p, t)P_2(t-1)\Gamma_2(p, t)]^{-1},$$

$$P_2(t) = [I_{mn} - L_4(t)\Gamma_2^T(p, t)]P_2(t-1),$$

$$W(p, t) = [\tilde{y}(t), \tilde{y}(t-1), \ldots, \tilde{y}(t-p+1)],$$

$$\Gamma_2(p, t) = [\hat{\psi}(t), \hat{\psi}(t-1), \ldots, \hat{\psi}(t-p+1)],$$

$$\hat{\psi}(t) = [-\tilde{w}_1(t-1), -\tilde{w}_1(t-2), \ldots, -\tilde{w}_1^T(t-n_i)].$$

The procedures of computing the parameter estimation vector $\hat{\theta}(t)$ by the M-D-MI-RGLS algorithm in Equations (33)–(47) are listed as follows:

1. Let $t = 1$, choose an innovation length $p$, set the initial values $\hat{\theta}(0) = 1_n/p_0$, $\hat{\theta}_n(0) = 1_{(mn) \times m}/p_0$, $P_1(0) = p_0I_n$, $P_2(0) = p_0I_{mn}$, $\hat{\theta}(t-i) = 1_m/p_0$ for $i = 1, 2, \ldots, n_i, p_0 = 10^6$.
2. Collect the observation data $\Phi_i(t)$ and $y(t)$, construct the information vector $\tilde{y}(t)$ using Equation (46).
3. Construct the stacked information matrices $\Gamma_1(p, t)$ and $\Gamma_2(p, t)$ using Equations (38) and (45).
4. Compute gain matrices $L_3(t)$ and $L_4(t)$ using Equations (35) and (42).
Figure 3. The flowchart of computing the M-D-MI-RGLS parameter estimates \( \hat{\theta}(t) \).

5. Compute covariance matrices \( P_1(t) \) and \( P_2(t) \) using Equations (36) and (43).

6. Compute \( \hat{y}_1(t) \) and \( \hat{w}(t) \) using Equations (39) and (47).

7. Construct the stacked output vector \( Y(p, t) \) using Equation (37) and the stacked noise output matrix \( W(p, t) \) using Equation (44).

8. Compute the large innovation vector \( E_1(p, t) \) and the innovation matrix \( E_2(p, t) \) using Equations (34) and (41).

9. Update the parameter estimates \( \hat{\theta}(t) \) and \( \hat{\theta}_u(t) \) using Equations (33) and (40).

10. Increase \( t \) by 1, go to Step 2.

The flowchart of computing \( \hat{\theta}(t) \) in the M-D-MI-RGLS algorithm is shown in Figure 3.

Obviously, we can obtain the M-D-RGLS algorithm when the innovation length \( p = 1 \). By expanding the innovation vectors \( e_1(t) \) and \( e_2(t) \) in the M-D-RGLS algorithm into a large innovation vector \( E_1(p, t) \) and an innovation matrix \( E_2(p, t) \) in the M-D-MI-RGLS algorithm, the data information of the system is used repeatedly, the accuracy of the parameter estimation is improved.

5. Example

Consider the following multivariate pseudo-linear autoregressive system:

\[
y(t) = \Phi(t)\theta + C^{-1}(z)v(t),
\]

\[
C(z) = I_2 + \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} z^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0.78 & -0.25 \\ 0.23 & -0.39 \end{bmatrix} z^{-1},
\]

\[
\theta = [\theta_1, \theta_2, \theta_3, \theta_4, c_{11}, c_{12}, c_{21}, c_{22}]^T = [2.28, 0.66, 1.14, 3.42]^T.
\]

The parameter vector to be estimated is

\[
\hat{\theta} = [\theta_1, \theta_2, \theta_3, \theta_4, c_{11}, c_{12}, c_{21}, c_{22}]^T = [2.28, 0.66, 1.14, 3.42, 0.78, -0.25, 0.23, -0.39]^T.
\]

| \( p \) | \( t \) | \( \theta_1 \) | \( \theta_2 \) | \( \theta_3 \) | \( \theta_4 \) | \( c_{11} \) | \( c_{12} \) | \( c_{21} \) | \( c_{22} \) | \( \delta (\%) \) |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 100 | 2.29864 | 0.65327 | 1.15998 | 3.37985 | 0.12221 | -0.29634 | 0.30171 | 0.41743 | 23.6837 |
| | 200 | 2.28186 | 0.66468 | 1.16771 | 3.40642 | 0.34208 | -0.24639 | 0.21832 | 0.31856 | 18.77530 |
| | 500 | 2.28077 | 0.66104 | 1.15388 | 3.42733 | 0.54506 | -0.20924 | 0.13124 | 0.08819 | 12.38976 |
| | 1000 | 2.27505 | 0.66276 | 1.14219 | 3.41717 | 0.64473 | -0.21133 | 0.12279 | 0.06144 | 8.45018 |
| | 2000 | 2.27249 | 0.66529 | 1.14387 | 3.41967 | 0.70977 | -0.23358 | 0.15898 | 0.19725 | 4.93468 |
| | 3000 | 2.27396 | 0.66371 | 1.14090 | 3.42247 | 0.71947 | -0.22704 | 0.16994 | 0.24929 | 3.76193 |
| 2 | 100 | 2.29864 | 0.65015 | 1.14985 | 3.38057 | 0.52688 | 0.14004 | 0.01507 | -0.29957 | 11.8232 |
| | 200 | 2.28060 | 0.65998 | 1.14831 | 3.40738 | 0.60500 | 0.07021 | 0.10721 | 0.26204 | 9.06153 |
| | 500 | 2.28024 | 0.66099 | 1.15405 | 3.42708 | 0.66760 | -0.03496 | 0.14836 | -0.33199 | 5.95427 |
| | 1000 | 2.27510 | 0.66266 | 1.14224 | 3.41710 | 0.71654 | -0.11304 | 0.17963 | -0.34902 | 3.71869 |
| | 2000 | 2.27244 | 0.66528 | 1.14392 | 3.41964 | 0.75085 | -0.18430 | 0.20677 | -0.37484 | 1.75917 |
| | 3000 | 2.27599 | 0.66368 | 1.14092 | 3.42247 | 0.74831 | -0.19210 | 0.20816 | -0.37930 | 1.59867 |
| 3 | 100 | 2.29849 | 0.64931 | 1.14907 | 3.38143 | 0.70161 | -0.00701 | 0.08034 | -0.13194 | 8.94898 |
| | 200 | 2.28350 | 0.65270 | 1.14981 | 3.40874 | 0.71611 | -0.05892 | 0.13921 | -0.15064 | 7.38634 |
| | 500 | 2.28055 | 0.66075 | 1.15410 | 3.42681 | 0.73435 | -0.12725 | 0.15798 | -0.25274 | 4.60845 |
| | 1000 | 2.27508 | 0.66258 | 1.14229 | 3.41712 | 0.75806 | -0.17421 | 0.18433 | -0.30212 | 2.87120 |
| | 2000 | 2.27241 | 0.66527 | 1.14381 | 3.41963 | 0.77484 | -0.22056 | 0.20912 | -0.34881 | 1.26608 |
| | 3000 | 2.27598 | 0.66367 | 1.14094 | 3.42249 | 0.76509 | -0.21768 | 0.20998 | -0.36107 | 1.14312 |

True values: 2.28000, 0.66000, 1.14000, 3.42000, 0.78000, -0.25000, 0.23000, -0.39000
In simulation, \( y(t) = \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} \in \mathbb{R}^2 \) is the output vector, \( \{ \Phi_s(t) \} \) is a 2 × 4 matrix sequence, \( v(t) = \begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix} \in \mathbb{R}^2 \) is a white noise vector with zero mean, \( \sigma_1^2 \) and \( \sigma_2^2 \) are the variance of \( v_1(t) \) and \( v_2(t) \). Taking the noise variances \( \sigma_1^2 = 0.20^2 \) and \( \sigma_2^2 = 0.30^2 \), using the M-D-RGLS algorithm (i.e. the M-D-MI-RGLS algorithm with \( p = 1 \)) and the M-D-MI-RGLS algorithm with \( p = 2 \) and \( p = 3 \) to estimate the parameters of this example system, we obtain the parameter estimates and their errors \( \delta := \| \hat{\theta}(t) - \theta \| / \| \theta \| \) shown in Table 1. The parameter estimation errors versus \( t \) are shown in Figure 4.

From Table 1 and Figure 4, we can draw the following conclusions:

1. The parameter estimation errors of the M-D-RGLS and the M-D-MI-RGLS algorithms become smaller with the data length \( t \) increases.
2. Under the same noise variances, the M-D-MI-RGLS algorithm has higher accurate parameter estimates than the M-D-RGLS algorithm.
3. Introducing the innovation length \( p \) can effectively improve the parameter estimation accuracy of the M-D-RGLS algorithm, as the innovation length \( p \) increases, the parameter estimates are getting more stationary.

6. Conclusions

This paper derives an M-D-RGLS algorithm and an M-D-MI-RGLS algorithm for identifying the multivariate pseudo-linear autoregressive system based on the decomposition technique and the multi-innovation identification theory. The simulation results indicate that the proposed algorithms are all effective and the M-D-MI-RGLS algorithm provide more accurate parameter estimation than the M-D-RGLS algorithm. The identification idea of the proposed method can be extended to study the parameter estimation problems of other scalar or multivariate, linear or nonlinear systems with coloured noises and applied to other fields.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the National Natural Science Foundation of China [grant number 61273194] and the 111 Project [grant number B12018].

Notes on contributors

Ping Ma received her bachelor degree from Jiangnan University, Wuxi, China in 2014. Currently, she is pursuing for master degree at the School of Internet of Things Engineering, Jiangnan University, Wuxi, China. Her research interests include system identification.

Feng Ding received his BSc degree from the Hubei University of Technology (Wuhan, China) in 1984, and his MSc and PhD degrees both from the Tsinghua University (Beijing, China) in 1991 and 1994, respectively. He has been a professor in the School of Internet of Things Engineering at the Jiangnan University (Wuxi, China) since 2004. His current research interests
include system identification and adaptive control. He authored five books on System Identification.

Quanmin Zhu is professor in control systems at the Department of Engineering Design and Mathematics, University of the West of England, Bristol, UK. He obtained his MSc in Harbin Institute of Technology, China in 1983 and PhD in Faculty of Engineering, University of Warwick, UK in 1989. His main research interest is in the area of nonlinear system modelling, identification, and control. He has published over 200 papers on these topics, edited five Springer books and one book for the other publisher, and provided consultancy to various industries. Currently Professor Zhu is acting as editor of International Journal of Modelling, Identification and Control, editor of International Journal of Computer Applications in Technology.

References

Anfinsen, H., Diagne, M., Aamo, O. M., & Krstic, M. (2017). Estimation of boundary parameters in general heterodirectional linear hyperbolic systems. Automatica, 79, 185–197.

Bessaoudi, T., Ben Hmida, F., & Hsieh, C. S. (2017). Robust state and fault estimation for linear descriptor stochastic systems with disturbances: A DC motor application. IET Control Theory and Applications, 11(5), 601–610.

Bravo, J. M., Suarez, A., Vasallo, M., & Alamo, T. (2016). Slide window bounded-error time-varying systems identification. IEEE Transactions on Automatic Control, 61(8), 2282–2287.

Chen, F. Y., Ding, F., Alsaedi, A., & Hayat, T. (2017). Data filtering based multi-innovation extended gradient method for controlled autoregressive autoregressive moving average systems using the maximum likelihood principle. Mathematics and Computers in Simulation, 132, 53–67.

Chen, M. T., Ding, F., Xu, L., Hayat, T., & Alsaedi, A. (2017). Iterative identification algorithms for bilinear-in-parameter systems with autoregressive moving average noise. Journal of the Franklin Institute, 354(17), 7885–7898.

Ding, F., Wang, X. H., Mao, L., & Xu, L. (2017). Joint state and multi-innovation parameter estimation for time-delay linear systems and its convergence based on the Kalman filtering. Digital Signal Processing, 62, 211–223.

Ding, F., Wang, F. F., Xu, L., Tasawar, H., & Ahmed, A. (2017). Parameter estimation for pseudo-linear systems using the auxiliary model and the decomposition technique. IET Control Theory and Applications, 11(3), 390–400.

Ding, F., Wang, F. F., Xu, L., & Wu, M. H. (2017). Decomposition based least squares iterative identification algorithm for multivariate pseudo-linear ARMA systems using the data filtering. Journal of the Franklin Institute, 354(3), 1321–1339.

Ding, F., Xu, L., & Zhu, Q. M. (2016). Performance analysis of the generalised projection identification for time-varying systems. IET Control Theory and Applications, 10(18), 2506–2514.

Escobar, J., & Poznyak, A. (2015). Benefits of variable structure techniques for parameter estimation in stochastic systems using least squares method and instrumental variables. International Journal of Adaptive Control and Signal Processing, 29(8), 1038–1054.

Hu, Y. B., Liu, B. L., & Zhou, Q. (2014). A multi-innovation generalized extended stochastic gradient algorithm for output nonlinear autoregressive moving average systems. Applied Mathematics and Computation, 247, 218–224.

Isaksson, A. J., Sjoberg, J., Tornqvist, D., Ljung, L., & Kok, M. (2015). Using horizon estimation and nonlinear optimization for grey-box identification. Journal of Process Control, 30, 69–79.

Li, M. H., & Liu, X. M. (2018). The least squares based iterative algorithms for parameter estimation of a bilinear system with autoregressive noise using the data filtering technique. Signal Processing, 147, 23–34.

Ma, P., & Ding, F. (2017). New gradient based identification methods for multivariate pseudo-linear systems using the multi-innovation and the data filtering. Journal of the Franklin Institute, 354(3), 1568–1583.

Ma, P., Ding, F., Alsaedi, A., & Hayat, T. (2018). Recursive least squares identification methods for multivariate pseudo-linear systems using the data filtering. Multidimensional Systems and Signal Processing, 29. doi:10.1007/s11045-017-0491-y

Ma, J. X., Xiong, W. L., Chen, J., & Ding, F. (2017). Hierarchical identification for multivariate Hammerstein systems by using the modified Kalman filter. IET Control Theory and Applications, 11(6), 857–869.

Mahnaz, H. (2017). Adaptive neural dynamic surface control of MIMO nonlinear time delay systems with time-varying actuator failures. International Journal of Adaptive Control and Signal Processing, 31(2), 275–296.

Mao, Y. W., & Ding, F. (2016). Parameter estimation for nonlinear systems by using the data filtering and the multi-innovation identification theory. International Journal of Computer Mathematics, 93(11), 1869–1885.

Meng, D. D. (2017). Recursive least squares and multi-innovation gradient estimation algorithms for bilinear stochastic systems. Circuits Systems and Signal Processing, 36(3), 1052–1065.

Milad, S., & Javad, A. (2017). Adaptive neural dynamic surface control of MIMO stochastic nonlinear systems with unknown control directions. International Journal of Adaptive Control and Signal Processing, 31(1), 97–121.

Pan, J., Jiang, X., Wan, X. K., & Ding, W. F. (2017). A filtering based multi-innovation extended stochastic gradient algorithm for multivariable control systems. International Journal of Control Automation and Systems, 15(3), 1189–1197.

Ramezani, Z., Arefi, M. M., Zargarzadeh, H., & Jahed-Motlagh, M. R. (2017). Neuro observer-based control of pure feedback MIMO systems with unknown control direction. IET Control Theory and Applications, 11(2), 213–224.

Sergey, D., & Svyatoslav, P. (2016). Constructive design of adaptive controllers for nonlinear MIMO systems with arbitrary switchings. IEEE Transactions on Automatic Control, 61(7), 2001–2007.

Shi, Z. W., Wang, Y., & Ji, Z. C. (2016a). Bias compensation based partially coupled recursive least squares identification algorithm with forgetting factors for MIMO systems: Application to PMSMs. Journal of the Franklin Institute, 353(13), 3057–3077.

Shi, Z. W., Wang, Y., & Ji, Z. C. (2016b). A multi-innovation recursive least squares algorithm with a forgetting factor for hammerstein CAR systems with backlash. Circuits Systems and Signal Processing, 35(12), 4271–4289.

Theodorakopoulos, A., & Rovithakis, G. A. (2016). Low-complexity prescribed performance control of uncertain
MIMO feedback linearizable systems. *IEEE Transactions on Automatic Control, 61*(7), 1946–1952.
Tohru, K., & Hajime, A. (2016). Linear approximation and identification of MIMO Wiener-Hammerstein systems. *Automatica, 71*, 118–124.
Wang, Y. J., & Ding, F. (2016). Recursive parameter estimation algorithms and convergence for a class of nonlinear systems with colored noise. *Circuits Systems and Signal Processing, 35*(10), 3461–3481.
Wang, D. Q., Mao, L., & Ding, F. (2017). Recasted models-based hierarchical extended stochastic gradient method for MIMO nonlinear systems. *IET Control Theory and Applications, 11*(4), 476–485.
Wang, D. Q., & Zhang, W. (2015). Improved least squares identification algorithm for multivariable Hammerstein systems. *Journal of the Franklin Institute, 352*(11), 5292–5307.
Xu, L. & Ding, F. (2017). Parameter estimation algorithms for dynamical response signals based on the multi-innovation theory and the hierarchical principle. *IET Signal Processing, 11*(2), 228–237.
Xu, L., & Ding, F. (2017). Recursive least squares and multi-innovation stochastic gradient parameter estimation methods for signal modeling. *Circuits Systems and Signal Processing, 36*(4), 1735–1753.
Zhang, X., Ding, F., Alsaeedi, A., & Hayat, T. (2017). Recursive parameter identification of the dynamical models for bilinear state space systems. *Nonlinear Dynamics, 89*(4), 2415–2429.
Zhu, Q. M., Wang, Y. J., Zhao, D. Y., Li, S. Y., & Billings, S. A. (2015). Review of rational (total) nonlinear dynamic system modelling, identification, and control. *International Journal of Systems Science, 46*(12), 2122–2133.
Zhu, Q. M., Yu, D. L., & Zhao, D. Y. (2017). An enhanced linear Kalman filter (EnLKF) algorithm for parameter estimation of nonlinear rational models. *International Journal of Systems Science, 48*(3), 451–461.