International Journal of Applied Mathematics in Control Engineering 2 (2019) 127-134

Contents lists available at YXpublications

International Journal of Applied Mathematics in Control Engineering

Journal homepage: www. YXpublications.com

Multi-innovation Least Squares for Identification of Wiener System based Decomposition Technique

Linwei Li*, Xianglong Liu

School of Electrical and Information Engineering, Zhengzhou University of Light Industry, Zhengzhou 450000, P.R. China

ARTICLE INFO

Article history:
Received 24 April 2019
Accepted 25 June 2019
Available online 28 December 2019

Keywords:
Time-delay
Sliding mode control
Nonlinear system
Non-affine
Neural networks

ABSTRACT

In this paper, a multi-innovation least-squares scheme is proposed for Wiener system on the basis of the decomposition technique. Firstly, the linear model is substituted into the special term of nonlinear model by using decomposition technique, an estimation model of parameter separation of linear and nonlinear model is established, which reduces the computation burden of algorithm. Secondly, a reference model is developed to handle the internal signal, which transforms the unknown internal signal into the indirectly measurable signal, which enhances the performance of identification method. Finally, the scalar innovation is revised to multi-innovation through the usage of certain length, which improves the accuracy of estimation scheme. Furthermore, the influences of different noise and different innovation length on the proposed algorithm are analyzed. The simulation results show that the proposed estimation approach outperforms the recursive least squares method in estimation accuracy and convergence rate.

1. Introduction

System identification technology is to select a model, standard and the collected data, and to fit the dynamic of the actual system. The model is selected based on the dynamic of the considered system, which is usually a simplified version of the real system. Standard includes mean square error, least squares standard, output error, etc. Data is the basis of system identification. In order to get accurate parameter estimation, the length of data should be large enough. To judge whether the system identification is successful, we can use model verification technology or some quantitative index values to test the performance of parameter estimation. In the past few decades，System identification technology has been widely used in control system, signal processing, fault diagnosis, event prediction and other fields, and has made a lot of achievements (Du et.al,2018, He et.al,2019,Yang et.al,2018).

System identification technology includes linear system identification technology and nonlinear system identification technology. The identification technology of linear system is suitable for dealing with linear system, and the performance of handling nonlinear system is not ideal. The linear system mainly involves ARX, AMRA, CARMA and gain system (Zhao et.al, 2017, Bedoui et.al,2015, Jin et.al,2014). The nonlinearity identification technology is used to address the nonlinear system, such as Hammerstein, Wiener, multivariable nonlinear system, neural network, nonlinear state space equation, T-S fuzzy model (Meng,2018, Gao et.al,2019, Zhao et.al,2019,Li et.al, 2018).

Wiener system is a modular non-linear system, which is composed of linear sub-models and memoryless non-linear modules in series. Such modular combination is very convenient and easy to use，and it is a kind of system that can describe the non-linear characteristics of many industrial processes (Lu, et al. 2019, Zhang et al. 2019, Kashima et al. 2018, Sadghi et al. 2019, Zhang et al. 2018). Therefore, the study of system identification of Wiener system has certain practical significance for the understanding of actual system modeling and working process. However, Wiener system is composed of different components in series, when the identification model is established, it leads to the phenomenon of coupling between non-linear parameters and linear parameters, i.e., over-parameter problem, redundant parameter estimation, or the non-uniqueness of estimated parameters, which increases the computation of identification algorithm, especially for the complex systems and large-scale systems. In addition, the internal signal of Wiener system cannot be measured directly, which also makes the difficulty of parameter identification.

Recursive least squares and its improved form are the most widely used identification algorithms in the field of system

* Corresponding author.
E-mail addresses: limengrui198801@163.com (L.W. Li)

identification. This is mainly due to its advantages of small calculation, strong applicability and convenient online estimation (Xia et al. 2019, Elisei-Iliescu et al. 2019, Kasai H. 2019). But it also has some shortcomings to be improved. For example, the performances of Recursive Least Square (RLS) for the identification of colored noise and strong noise are not ideal. The low utilization of information in the past and current time of the system results in low accuracy, and it is powerless for time-varying systems. Therefore, in order to solve these problems, a large number of improved algorithms have been designed by researchers and successfully applied in the field of system identification. Wang et al. (Wang et al. 2018) calculated the inverse function of piecewise affine nonlinear function in Hammerstein system, then established a parameterized linear regression identification model, and used the RLS to estimate the parameters of the model under bounded noise. Shi et al. (Shi et al. 2018) used the linear dynamic model to approximate the motor according to the working principle of the motor. To improve the accuracy of RLS, the discount factor is used to modify the corrected gain of RLS and apply it to the parameter estimation of the motor model. Ding et al. (Ding et al. 2018) explored the hierarchical principle to divide the system model into several sub-models identification such that the computational complexity of the system is decreased, and then proposed a least squares identification algorithm based on hierarchical method to estimate the system parameters. Hafezi et al. (Hafezi et al. 2019) proposed a recursive generalized least squares scheme to solve the problem of ARMA noise model, and compared it with RML method to verify the effectiveness of the proposed algorithm. Wei et al. (Wei et al. 2019) established the equivalent circuit model of the battery system, combined the multi-innovation theory with least squares to form a multi-innovation least squares identification method, and can effectively estimate of battery parameters under battery charging and discharging experiments. In Gan et al. (Gan et al. 2019), the projection algorithm is used to estimate the nonlinear parameters of the system, and then the multi-innovation least squares estimation scheme is presented to estimate the parameters of the linear part. Compared with some existing estimation methods, the proposed algorithm performs better. From the above literature, we can see that there are many ways to improve the least squares, but the improvement of multi-innovation theory is a relatively new identification method in recent years. At present, the application of multi-innovation least squares in linear systems is more, and the non-linear systems are relatively less. It is more difficult for a non-linear system such as Wiener, which exists the output signal distortion because the non-linear part is behind the system.

In this paper, a decomposition-based multi-innovation Least Squares method (D-MILS) is proposed to solve the over-parameter problem and the non-unique parameter problem of Wiener system. The decomposition technique is used to construct the identification model of the system with each parameter separated from each other, which reduces the computational burden of the identification algorithm. The reference model is established to solve the problem of unknown internal signals, and to solve the immeasurable internal signals. *p* data are used to modify the utilization rate of observation vectors, to improve the estimation accuracy and convergence speed. Finally, the effectiveness and advantages of the proposed estimation scheme are compared.

2. Problem Description

Consider the Wiener system as shown in Fig.1

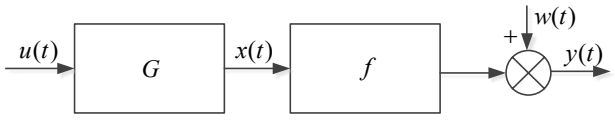

Fig. 1. Block diagram of Wiener system

where G, f denote the linear part and nonlinear part. $u(t), y(t)$ are the input-output of system. $x(t), w(t)$ are inner and noise signals.

Wiener system can be represented by the following equation

$$x(t) = G(z)u(t) \tag{1}$$

$$y(t) = f(x(t)) + w(t) = \sum_{i=1}^{n} \lambda_i [x(t)]^i + w(t) \tag{2}$$

$$G(z) = \frac{B(z)}{A(z)} \tag{3}$$

where $f(x(t))$ is a combination of basis functions with finite order. $G(z)$ denotes rational fraction, where $B(z) = b_1 z^{-1} + \cdots + b_n z^{-n}$, $A(z) = 1 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_n z^{-n}$.

When the Eqs.(1) and (3) are substituted into (2), we have

$$y(t) = \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \frac{b_j}{a_j} [u(t-j)]^i + w(t)$$

which results in the block parameters and too many parameters need to be estimated, which increases the computational burden of the algorithm, and makes it impossible to obtain the unique identification model because there are numerous solutions for the composite parameters as long as the results are correct.

In order to avoid this problem, the decomposition technology (Ding et al. 2016, Voros.J.2019) is used to construct the identification model in which system parameters are separated from each other, so as to reduce the calculation and improve the operation efficiency.

To achieve this goal, based on Eq.(2), $\lambda_1 x(t)$ is selected as the key-term, the other items remain unchanged, we have

$$\begin{aligned} y(t) = & -\lambda_1 a_1 x(t-1) - \lambda_1 a_2 x(t-2) -, \cdots, -\lambda_1 a_n x(t-n) \\ & + \lambda_1 b_1 u(t-1) + \lambda_1 b_2 u(t-2) +, \cdots, +\lambda_1 b_n u(t-n) \\ & + \lambda_2 x(t)^2 + \lambda_3 x(t)^3 +, \cdots, +\lambda_n x(t)^n + w(t) \end{aligned} \tag{4}$$

Eq.(4) is written in compact form as follows:

$$y(t) = \varphi^T(t)\theta + w(t) \tag{5}$$

$$\begin{aligned} \varphi(t) = & [-x(t-1), -x(t-2), \cdots, -x(t-n), \\ & u(t-1), u(t-2), \cdots, u(t-n), x(t)^2 \\ & , x(t)^3, \cdots, x(t)^n]^T \end{aligned} \tag{6}$$

$$\begin{aligned} \theta = & [\lambda_1 a_1, \lambda_1 a_2, \cdots, \lambda_1 a_n, \lambda_1 b_1, \lambda_1 b_2, \cdots, \\ & \lambda_1 b_n, \lambda_2, \lambda_3, \cdots, \lambda_n]^T \end{aligned} \tag{7}$$

Before implementing the identification algorithm, some

appropriate assumptions have been made for parameter identification. These assumptions have been widely used in the field of parameter identification and system modeling (Li et al. 2009, Wahlberg et al. 2018)

Assumption : (1) The input is a persistent excitation signal, which makes the system identifiable; (2) the linear subsystem is stable; (3) to obtain a unique identification model, λ_1=1 is given.

Assumption (1) shows that the all mode of system can be excited, and the system is identifiable. In Assumption (2), there is no pole zero cancellation in the numerator and denominator of a linear system, which is the basis of system identification. Assumption (3) indicates that a multiple of the difference between the estimated parameter and the actual parameter, according to λ_1=1 , we can obtain the independent parameters.

Remark 1: According to the Eq.(7) and Assumption (3), we know that the linear and non-linear parameters in the estimated parameters are separated from each other and do not contain block parameters based on the decomposition technology, which reduces the computational cost and improves the identification performance.

3. Estimation scheme

The emergence of RLS is to improve the identification ability of least squares and adapt to online identification. Based on the Eq.(5), the cost function $J(\theta)=\left\|y(t)-\varphi^T(t)\theta\right\|^2$ is built.

Define $\begin{cases} y(1)=\varphi^T(1)\theta+w(1) \\ y(2)=\varphi^T(2)\theta+w(2) \\ \vdots \\ y(t)=\varphi^T(t)\theta+w(t) \end{cases}$ and the matrix of the above equation are written as

$$\begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(t) \end{bmatrix}=\begin{bmatrix} \varphi^T(1) \\ \varphi^T(2) \\ \vdots \\ \varphi^T(t) \end{bmatrix}\theta+\begin{bmatrix} w(1) \\ w(2) \\ \vdots \\ w(t) \end{bmatrix}$$

The above equation can be rewritten as follows:

$$\boldsymbol{Y}_t=\boldsymbol{H}_t\theta+\boldsymbol{W}_t$$

where

$$\boldsymbol{Y}_t=\begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(t) \end{bmatrix},\ \boldsymbol{H}_t=\begin{bmatrix} \varphi^T(1) \\ \varphi^T(2) \\ \vdots \\ \varphi^T(t) \end{bmatrix},\ \boldsymbol{W}_t=\begin{bmatrix} w(1) \\ w(2) \\ \vdots \\ w(t) \end{bmatrix}$$

According to $J(\theta)=\left\|y(t)-\varphi^T(t)\theta\right\|^2$, we have

$$\begin{aligned} J(\theta)&=\sum_{j=1}^{t}\left[y(j)-\varphi^T(j)\theta\right]^2 \\ &=W_t^TW_t \\ &=\left(Y_t-H_t\theta\right)^T\left(Y_t-H_t\theta\right) \\ &=\left\|Y_t-H_t\theta\right\|^2 \end{aligned}$$

According to prediction error method, by calculating the derivative of the object function J to the parameter vector θ, we have

$$\frac{\partial J(\theta)}{\theta}=-2\boldsymbol{H}^T{}_t(\boldsymbol{Y}_t-\boldsymbol{H}_t\theta)\,|_{\theta=\hat{\theta}}=0$$

or

$$\hat{\theta}=(\boldsymbol{H}^T{}_t\boldsymbol{H}_t)^{-1}\boldsymbol{H}^T{}_t\boldsymbol{Y}_t$$

Before deriving recursive least squares, we first introduce a lemma about matrix inverse operation

Lemma 1: Matrix inversion lemma: assume that A,B,C are matrices, then $(A+BC)^{-1}=A^{-1}B(I+CA^{-1}B)^{-1}CA^{-1}$ holds.

Proof: If $AB=I$, then we have $A^{-1}=B$, vice versa.

$$\begin{aligned} &(A+BC)[A^{-1}B(I+CA^{-1}B)^{-1}CA^{-1}] \\ &=I-B(I+CA^{-1}B)^{-1}CA^{-1}+BCA^{-1}-BCA^{-1}B(I+CA^{-1}B)^{-1}CA^{-1} \\ &=I+BCA^{-1}-B(I+CA^{-1}B)^{-1}CA^{-1}-BCA^{-1}B(I+CA^{-1}B)^{-1}CA^{-1} \\ &=I+BCA^{-1}-B(I+CA^{-1}B)(I+CA^{-1}B)^{-1}CA^{-1} \\ &=I+BCA^{-1}-BICA^{-1} \\ &=I \end{aligned}$$

For the same reason, we have

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

This is proof of lemma one.

Define

$$P^{-1}(t)=\sum_{j=1}^{t}\varphi(j)\varphi^T(j)=P^{-1}(t-1)+\varphi(t)\varphi(t)^t,$$

$$\varepsilon(t)=\sum_{j=1}^{t}\varphi(j)y(j)=\varepsilon(t-1)+\varphi(t)y(t),\varepsilon(0)=0.$$

then, $\hat{\theta}=(\boldsymbol{H}^T{}_t\boldsymbol{H}_t)^{-1}\boldsymbol{H}^T{}_t\boldsymbol{Y}_t$ can be written as

$$\hat{\theta}=P(t)\varepsilon(t)$$

By using the Lemma 1, combining with matrix inversion, we have

$$P(t)=P(t-1)-\frac{P(t-1)\varphi(t)\varphi^T(t)P(t-1)}{1+\varphi^T(t)P(t-1)\varphi(t)},P(0)=p_0I$$

For the estimation parameter, we have

$$\begin{aligned} \hat{\theta}&=P(t)\varepsilon(t) \\ &=P(t)\begin{bmatrix} H_{t-1} \\ \varphi^T(t) \end{bmatrix}^T\begin{bmatrix} Y_{t-1} \\ y(t) \end{bmatrix} \\ &=P(t)\begin{bmatrix} H^T{}_{t-1} & \varphi(t) \end{bmatrix}\begin{bmatrix} Y_{t-1} \\ y(t) \end{bmatrix} \end{aligned}$$

$$
\begin{aligned}
&= P(t)[H^{T}{}_{t-1}Y_{t-1} + \varphi(t)y(t)] \\
&= P(t)[P^{-1}(t-1)P(t-1)H^{T}{}_{t-1}Y_{t-1} + \varphi(t)y(t)] \\
&= P(t)[P^{-1}(t-1)\hat{\theta}(t-1) + \varphi(t)y(t)] \\
&= P(t)[P^{-1}(t) - \varphi(t)\varphi^{T}(t)]\hat{\theta}(t-1) + P(t)\varphi(t)y(t) \\
&= \hat{\theta}(t-1) + P(t)\varphi(t)[y(t) - \varphi^{T}(t)\hat{\theta}(t-1)]
\end{aligned}
$$

Further, the expression of RLS can be obtained as follows:

$$\hat{\theta}(t) = \hat{\theta}(t-1) + P(t)\varphi(t)[y(t) - \varphi^{T}(t)\hat{\theta}(t-1)] \tag{8}$$

$$e(t) = y(t) - \varphi^{T}(t)\hat{\theta}(t-1) \tag{9}$$

$$P(t) = P(t-1) - \frac{P(t-1)\varphi(t)\varphi^{T}(t)P(t-1)}{1+\varphi^{T}(t)P(t-1)\varphi(t)}, P(0) = p_0 I \tag{10}$$

where $P(t)$ is the covariance matrix. $e(t)$ represents the scalar innovation.

The form of gain vector can be written as follows:

$$
\begin{aligned}
L(t) &= P(t)\varphi(t) \\
&= P(t-1)\varphi(t) - \frac{P(t-1)\varphi(t)\varphi^{T}(t)P(t-1)\varphi(t)}{1+\varphi^{T}(t)P(t-1)\varphi(t)} \\
&= P(t-1)\varphi(t)[1 - \frac{P(t-1)\varphi(t)\varphi^{T}(t)P(t-1)\varphi(t)}{1+\varphi^{T}(t)P(t-1)\varphi(t)}] \\
&= \frac{P(t-1)\varphi(t)}{1+\varphi^{T}(t)P(t-1)\varphi(t)}
\end{aligned}
$$

Eqs.(8) - (10) shows that the parameter estimation is conducted which only use current time data information, but not consider the other time information. This leads to the unsatisfactory estimation ability of RLS in some cases. In order to make full use of the system information, it is necessary to modify the scalar innovation. The most appropriate way is to expand the scalar innovation through several groups of data, then, the current and past time data are used in parameter updating process, which improves the utilization rate of data and thus increases the performance of identification.

We use p data to modify the scalar innovation, it yields multi-innovation $E(p,t)$, which has the following form:

$$E(p,t) = y(p,t) - \phi^{T}(p,t)\hat{\theta}(t-1). \tag{11}$$

Then, the other variables can be written as:

$$\phi(p,t) = [\varphi(t), \varphi(t-1), \cdots, \varphi(t-p+1)], \tag{12}$$

$$y(p,t) = [y(t), y(t-1), \cdots, y(t-p+1)]^{T}. \tag{13}$$

According to the modification of Eqs.(11) - (13), the expression of multi-innovation least squares is listed as follows.

$$\hat{\theta}(t) = \hat{\theta}(t-1) + P(t)\phi(p,t)E(p,t), \tag{14}$$

$$E(p,t) = y(p,t) - \phi^{T}(p,t)\hat{\theta}(t-1), \tag{15}$$

$$P^{-1}(t) = P^{-1}(t-1) + \phi(p,t)\phi^{T}(p,t), P(0) = p_0 I. \tag{16}$$

In (16), the inverse of the covariance matrix needs to be computed, which produces a high computational effort. In order to solve this problem, the principle of matrix inversion is adopted. Eqs. (14) - (16) can be rewritten as follows:

$$\hat{\theta}(t) = \hat{\theta}(t-1) + L(t)[y(p,t) - \phi^{T}(p,t)\hat{\theta}(t-1)] \tag{17}$$

$$L(t) = P(t-1)\phi(p,t)/[I - \phi^{T}(p,t)P(t-1)\phi(p,t)] \tag{18}$$

$$P(t) = P(t-1) + L(t)\phi^{T}(p,t)P(t-1), P(0) = p_0 I \tag{19}$$

Remark 2: By expanding the scalar innovation into multi-innovation vector, the system data are further utilized compared with the scalar innovation. Multi-innovation ways can lift the data utilization and enhance the precision and speed of the estimator.

Although the identification algorithm is designed, the estimation scheme (17)-(19) cannot be implemented because the data vector contains internal signal $x(t)$. To deal with this issue, we design the reference model according to the idea of reference model identification (Li et al. 2018, Liu et al. 2019), and use the output of the reference model to replace the internal signal, as shown in Fig.2.

The expression of the reference model $x_{ref}(t)$ is as follows:

$$x_{ref}(t) = -\sum_{i=1}^{n}\hat{a}_i x_{ref}(t-i) + \sum_{j=1}^{n}\hat{b}_j u(t-j) \tag{20}$$

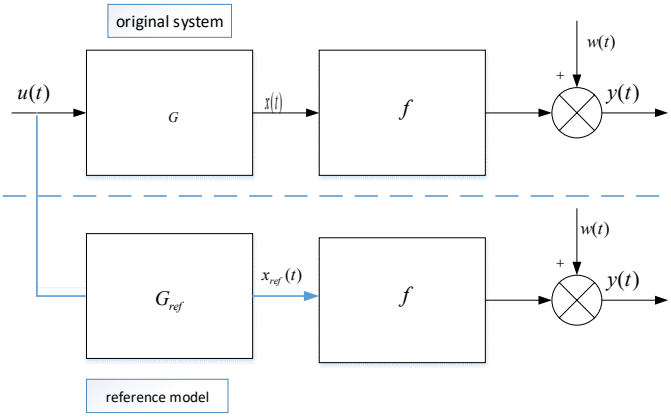

Fig.2. The system based on reference model

The internal signal $x(t)$ is replaced by the output $x_{ref}(t)$, the $\varphi(t)$ is also replace by $\varphi_{\text{ref}}(t)$. The data matrix can be written as

$$\phi_{ref}(p,t) = [\varphi_{ref}(t), \varphi_{ref}(t-1), \cdots, \varphi_{ref}(t-p+1)]$$

the multi-innovation is expressed as

$$E(p,t) = y(p,t) - \phi_{ref}{}^{T}(p,t)\hat{\theta}(t-1)$$

The multi-innovation least squares can be rewritten based reference model as follows:

$$\hat{\theta}(t) = \hat{\theta}(t-1) + L(t)[y(p,t) - \phi_{\text{ref}}{}^{T}(p,t)\hat{\theta}(t-1)] \tag{21}$$

$$L(t) = P(t-1)\phi_{\text{ref}}(p,t)/[I - \phi_{\text{ref}}{}^{T}(p,t)P(t-1)\phi_{\text{ref}}(p,t)] \tag{22}$$

$$P(t) = P(t-1) + L(t)\phi_{\text{ref}}{}^{T}(p,t)P(t-1), P(0) = p_0 I \quad (23)$$

To summarize, we list the steps involved in the MILS algorithm to recursively compute the parameter estimation vector $\hat{\theta}(t)$ as t increases:

Step 1: Collect the input\output data $y(t), \varphi(t)$ and choose the length of sample data N.

Step 2: For initialize, Let $t = 1, \quad p_0 = 10^6$, $\hat{\theta}(0) = I / p_0$.

Step 3: Form $\phi_{\text{ref}}(p,t)$ by (12), and $Y(p,t)$ by (13).

Step 4: Compute $L(t)$ by (22), $P(t)$ by (23), $\hat{\theta}(t)$ by (21).

Step 5: When $t = N+1$, then terminate the procedure and obtain the estimates of the parameter vector; otherwise, increment t by 1 and go to step 3.

The convergence of MILS is summarized as follows:

Let us introduce some notation first. E denotes the expectation operator, the norm of the matrix X is defined by $\|X\|^2 = tr[XX^T]$

$|X| = \det[X]$ represents the determinant of the square matrix X; $\lambda_{\max}[X], \lambda_{\min}[X]$ represent the maximum and minimum eigenvalues of X, respectively.

Next, we establish the convergence properties of the MILS algorithm.

Lemma 2: For the MILS and any the innovation length, then the following inequality holds:

$$\sum_{t=1}^{\infty} \frac{\varphi^T(t-i)P(t)\varphi(t-i)}{[\ln]|P^{-1}(t)|]^c} < \infty, c > 1, \quad i = 1, \cdots, p$$

Proof: From the definition of $P(t)$, we have

$$\begin{aligned} P^{-1}(t-1) &= P^{-1}(t) - \phi(p,t)\phi^T(p,t) \\ &\le P^{-1}(t) - \varphi(t-i)\varphi^T(t-i) \\ &= P^{-1}(t)[I - P(t)\varphi(t-i)\varphi^T(t-i)] \end{aligned}$$

Taking the determinant of both sides of the above equation and using det[I + DE] = det[I, + ED] give

$$\begin{aligned} \left|P^{-1}(t-1)\right| &\le \left|P^{-1}(t)\right| \left|I - P(t)\varphi(t-i)\varphi^T(t-i)\right| \\ &= \left|P^{-1}(t)\right| \left|1 - \varphi^T(t-i)P(t)\varphi(t-i)\right| \end{aligned}$$

Thus,

$$\left|\varphi^T(t-i)P(t)\varphi(t-i)\right| \le \frac{\left|P^{-1}(t)\right| - \left|P^{-1}(t-1)\right|}{\left|P^{-1}(t)\right|}$$

By dividing the $[\ln|P^{-1}(t)|]^c$, and summing for *t* give

$$\sum_{t=1}^{\infty} \frac{\left|\varphi^T(t-i)P(t)\varphi(t-i)\right|}{[\ln|P^{-1}(t)|]^c} \le \sum_{t=1}^{\infty} \frac{\left|P^{-1}(t)\right| - \left|P^{-1}(t-1)\right|}{\left|P^{-1}(t)\right|[\ln|P^{-1}(t)|]^c}$$

$$\le \int_{|P^{-1}(0)|}^{|P^{-1}(\infty)|} \frac{dx}{x[\ln x]^c} = \frac{-1}{c-1}\left[\frac{1}{[\ln x]^{c-1}}\right]_{|P^{-1}(0)|}^{|P^{-1}(\infty)|}$$

$$= \frac{1}{c-1}\left[\frac{1}{[\ln|P^{-1}(0)|]^{c-1}} - \frac{1}{[\ln|P^{-1}(\infty)|]^{c-1}}\right] < \infty$$

This completes the proof of Lemma 1.

Theorem 1: For the considered system and MILS algorithm, assume that $\{w(t), \mathrm{F}(t)\}$ is a martingale difference sequence, where $\{\mathrm{F}(t)\}$ is a algebra sequence generated by the observation data up to and including time *t*. The sequence $\{w(t)\}$ satisfies the noise assumptions:

(1) $E[w(t) \mid \mathrm{F}(t-1)] = 0$

(2) $E[w^2(t) \mid \mathrm{F}(t\text{-}1)] = \sigma^2{}_w(t) < \sigma^2 < \infty$

and there exist $\alpha, \ \beta > 0, \alpha_0 \ge 0$ such that the following generalized persistent excitation condition holds:

(3) $\alpha I \le \frac{1}{t}\sum_{i=1}^{t} \phi(i)\phi^T(i) \le \beta t^{\alpha_0} I$

Then the parameter estimation error almost surely converges to zero, i.e.,

$$\lim_{t\to\infty} \left\|\hat{\theta}(t) - \theta\right\| = 0$$

From this theorem, we can see that the parameter estimates given by the MILS algorithm always converge to their true values and are more accurate than those obtained by the standard LS algorithm for the same data length, see the example later.

The goal of this paper is to design a low computational complexity and high accuracy identification algorithm based on the collected input-output data, to estimate the parameters of the system, and to analyze the advantages of D-MILS by comparing the decomposition technique of recursive least squares method (D-RLS).

Remark 3: By using Eq.(20), the unknown inner signal is converted into indirectly measurable signals according to the reference model of original signal, which makes the unmeasurable signal close to the actual signal and improves the identification nature.

4. Example

Consider the Wiener system:

$$x(t) = -0.65x(t-1) - 0.4x(t-2) + 0.5u(t-1) + 0.25u(t-2)$$

$$y(t) = x(t) + 0.35x^2(t) + 0.5x^3(t) + w(t)$$

Based on the considered system, we know that real parameters are $a_1 = 0.5, a_2 = 0.25, b_1 = 0.65, b_2 = 0.4, \lambda_2 = 0.35, \lambda_3 = 0.5$. The input signal chooses a random signal whose mean value is zero and variance is one, and the noise is white noise. The estimated values of D-RLS and D-MILS under different sample sizes are listed as Tables 1-2. From the Tabs.1-2, we know that with the addition of sample, the estimated parameters fluctuate near the real values, and the closer to the expected value, the smaller the fluctuation amplitude. Although both algorithms can estimate the parameters of Wiener system, but the estimated values of the proposed algorithm D-MILS are closer to the real values with high

accuracy compared to D-RLS.

Tab.1. The identification results of D-RLS

N	100	500	800	1000	2000	True value
a_1	0.667	0.500	0.500	0.500	0.501	0.500
a_2	0.271	0.252	0.251	0.250	0.249	0.250
b_1	0.709	0.654	0.652	0.650	0.647	0.650
b_2	0.614	0.399	0.401	0.399	0.401	0.400
λ_2	0.340	0.348	0.349	0.350	0.350	0.350
λ_3	0.081	0.498	0.496	0.498	0.498	0.500

Tab.2. The estimation results of D-MILS

N	100	500	800	1000	2000	Real value
a_1	0.503	0.501	0.503	0.503	0.501	0.500
a_2	-0.061	0.133	0.169	0.184	0.214	0.250
b_1	0.035	0.428	0.499	0.527	0.582	0.650
b_2	0.228	0.345	0.364	0.372	0.383	0.400
λ_2	0.440	0.374	0.369	0.364	0.356	0.350
λ_3	0.497	0.508	0.498	0.496	0.501	0.500

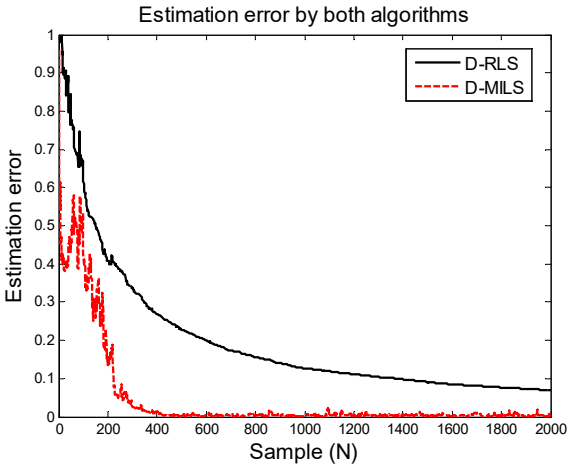

Fig.3. The estimation error by both methods

Fig.3 shows the results of D-RLS and D-MILS estimation errors. It can be seen from the graph that with the increase of samples, the estimation errors of the two estimation schemes are gradually reduced, a small steady-state value is finally reached, which shows that both D-RLS and D-MILS can identify the parameters of Wiener system. Fig. 3 also depicts that the proposed algorithm has better estimation accuracy and faster convergence speed than D-RLS. Fig. 3 is an identification performance of the proposed algorithm under different innovation lengths. It can be seen from Fig.4 that with the increase of innovation lengths, the convergence rate of estimation error is faster, but the oscillation is more serious. The reason for this phenomenon is that although the identification algorithm uses more and more system information to improve the utilization rate of identification data, noise information is also further utilized at the same time. At this time, the impact of noise on identification performance is also increasing, which leads to the deterioration of estimation performance.

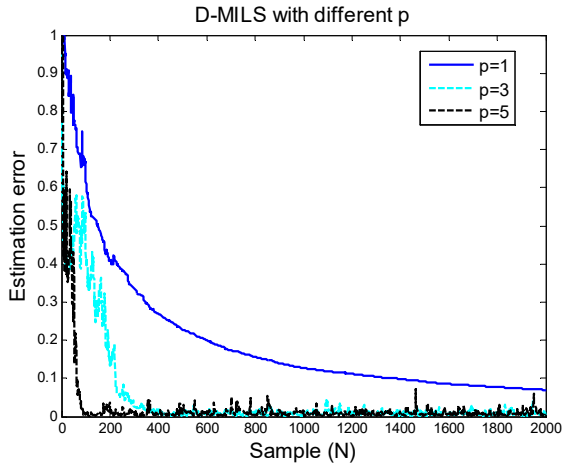

Fig.4. The estimation error with different innovation length

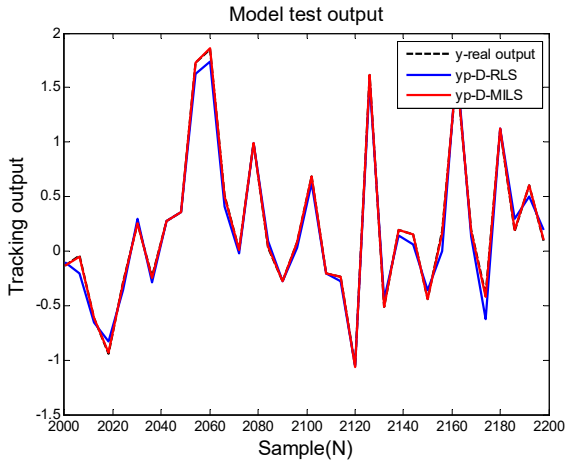

Fig.5. Model test output

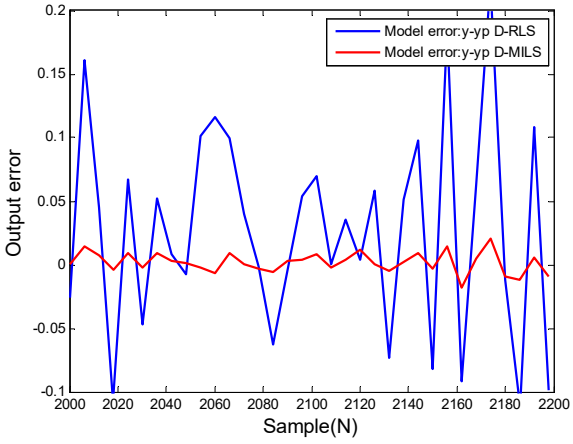

Fig.6. Output error

One of the methods to evaluate effectiveness of the parameter estimation is to build the corresponding estimation model based on the estimated parameters, and compare the difference between the model output and the actual output. When the difference is small, it shows that the estimation model can describe the output of the actual system and the identification algorithm can accurately estimate the parameters of the system. However, when the

difference is large, it shows that the estimation algorithm cannot effectively estimate the parameters of the system, and it can be considered that the parameter estimation is the failure. Figs.5-6 show the model validation and output error results. It can be observed that the model output can track the real output, but the presented method provides a higher tracking result than the D-RLS algorithm.

Table 3 displays the identification error results of the proposed algorithm under different noise levels. From the Tab.3, it can be seen that the parameter estimates are close to the real parameters under weak noise, and the estimates fluctuate greatly near the real values under strong noise, which also shows that noise has some adverse effects on the identification algorithm.

σ^2	N	100	500	800	1000	2000	True value
$(0.1)^2$	a_1	0.667	0.500	0.500	0.500	0.501	0.500
	a_2	0.271	0.252	0.251	0.250	0.249	0.250
	b_1	0.709	0.654	0.652	0.650	0.647	0.650
	b_2	0.614	0.399	0.401	0.399	0.401	0.400
	λ_2	0.340	0.348	0.349	0.350	0.350	0.350
	λ_3	0.081	0.498	0.496	0.498	0.498	0.500
$(0.5)^2$	a_1	0.513	0.496	0.492	0.507	0.483	0.500
	a_2	0.242	0.250	0.258	0.263	0.217	0.250
	b_1	0.653	0.649	0.672	0.682	0.562	0.650
	b_2	0.403	0.396	0.416	0.397	0.386	0.400
	λ_2	0.366	0.343	0.353	0.354	0.317	0.350
	λ_3	0.482	0.507	0.504	0.478	0.559	0.500

Tab.3. The performance of D-MILS with different noise

5. Summary

Aiming at the parameter identification of nonlinear Wiener system, the decomposition technique is used to select the decomposition term and substitute it into the corresponding expression, and to construct the identification model with fewer parameters in the estimated parameter vector, which improves the efficiency of the algorithm. For the internal variables of the system, the output of the reference model is used to replace the unknown internal variables, which lifts the identification performance. In order to raise the convergence speed and estimation accuracy of the recursive least squares, some data are used to modify the innovation vector so that the scalar innovation can be transformed into a multi-innovation, and the multi-innovation least squares method can be obtained. Finally, the performances of the proposed algorithm under different noise and different length of innovation are analyzed by simulation. The results show that the proposed algorithm has certain advantages.

Acknowledgements

This work is supported by Shandong Natural Science Foundation of China(ZR2017MF048), Scientific Research Foundation of Shandong University of Science and Technology for Recruited Talents (2016RCJJ035), Tai'an Science and Technology development program (2017GX0017)

References

Du W, Zhao X., Du H. et.al,2018, Fault Diagnosis of Analog Circuit based on Extreme Learning Machine, International Journal of Applied Mathematics in Control Engineering ,1(1) 9-14

He Y. Cheng L, Liu J,2019, Research and Design of Air-water Amphibious Six-rotor UAV Based on the New Air Structure, International Journal of Applied Mathematics in Control Engineering,2(1),2019:51-57

Yang Z. ,Fu J. ,Wang L. et.al. Back-stepping Sliding Mode with Unidirectional Auxiliary Surfaces for HSV with Attitude Constrains, International Journal of Applied Mathematics in Control Engineering, 1(2),2018, 187-193

Zhao, Y., Fatehi, A., & Huang, B. (2017). A data-driven hybrid arx and markov chain modeling approach to process identification with time-varying time delays. IEEE Transactions on Industrial Electronics, 64(5), 4226-4236.

Bedoui, S., & Abderrahim, K. (2015). Armax time delay systems identification based on least square approach. Ifac Papersonline, 48(28), 1100-1105.

Jin, G., & Zhao, Y. L. (2014). Identification of the gain system with quantized observations and bounded persistent excitations. Science China(Information Sciences), 57(1), 12205-012205.

Gao X.,Zhang A., Wang S. et.al., 2019,Adaptive High Order Neural Network Identification and Control for Hysteresis Nonlinear System with Bouc-Wen Model, International Journal of Applied Mathematics in Control Engineering,2(1),2019:17-24

Meang R.X., 2018. Adaptive Parameter Estimation for Multivariable Nonlinear CARMA Systems, International Journal of Applied Mathematics in Control Engineering, 1(1), 96-102

Li, C., Wen, Z., Nan, Z., & Lai, X. (2018). An evolving t–s fuzzy model identification approach based on a special membership function and its application on pump-turbine governing system. Engineering Applications of Artificial Intelligence, 69, 93-103.

Xiaodong Zhao, Hui Du, Research on Human Walking Feature Extraction and Identification Recognition System, International Journal of Applied Mathematics in Control Engineering,2(1),2019:81-86

Lu M, Wang H, Lin J, et al.2019. A nonlinear Wiener system identification based on improved adaptive step-size glowworm swarm optimization algorithm for three-dimensional elliptical vibration cutting. The International Journal of Advanced Manufacturing Technology，103(5), 2865-2877.

Zhang Y., Jiang N. Wang L.P.,2019. Reliability Analysis of Minimum Samples of NC Turntable Based on Wiener Process Journal of Tsinghua University (Natural Science Edition) 59(02),91-95.

Kashima K, Aoyama H, Ohta Y. 2018.Stable Process Approach to Analysis of Systems Under Heavy-Tailed Noise: Modeling and Stochastic Linearization. IEEE Transactions on Automatic Control,64(4),1344-1357.

Sadeghi M, Farrokhi M. 2019.Real-time identification of nonlinear multiple-input-multiple-output systems with unknown input time delay using Wiener model with Neuro-Laguerre structure. International Journal of Adaptive Control and Signal Processing, 33(1),157-174.

Zhang M., Liu J., Liao L. et.al. 2018.Residual Life Assessment Method for Oil-paper Insulation Aging of Transformer Based on Wiener Model. Journal of Electrical Technology, 33(21),5098-5108.

Xia H, Yang Y, Ding F, et al. 2019.Maximum likelihood-based recursive least-squares estimation for multivariable systems using the data filtering technique. International Journal of Systems Science, 50(6),1121-1135.

Elisei-Iliescu C, Paleologu C, Benesty J, et al. 2019.Recursive Least-Squares Algorithms for the Identification of Low-Rank System. IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP), 27(5),903-918.

Kasai H.2019. Fast online low-rank tensor subspace tracking by CP decomposition using recursive least squares from incomplete observations. Neurocomputing, 347, 177-190.

Wang J., Dao W. 2018.Recursive least squares identification for piecewise affine Hammerstein models. International Journal of Intelligent Computing and Cybernetics,11(2),234-153

Shi J., Ge B., Lv Y., et al. 2018.Research on On-line Parameter Identification Method for Permanent Magnet Synchronous Motor.Journal of Motor and Control, 22(03),17-24.

Ding F., Chen H., Xu L., et al.2018. A hierarchical least squares identification algorithm for Hammerstein nonlinear systems using the key term separation. Journal of the Franklin Institute, 355(8), 3737-3752.

Hafezi Z, Arefi M. 2019.Recursive generalized extended least squares and RML algorithms for identification of bilinear systems with ARMA noise. ISA transactions, 88, 50-61.

Wei Z., Yuan K. Cheng L., et al. 2019.Lithium battery parameter identification based on multi-innovation least squares algorithm. Power system automation, 43(15),139-145

Gan M, Chen X X, Ding F, et al.2019. Adaptive RBF-AR Models Based on Multi-Innovation Least Squares Method. IEEE Signal Processing Letters, 26(8), 1182-1186.

Ding F., Wang F., Xu L., et al. 2016.Parameter estimation for pseudo-linear systems using the auxiliary model and the decomposition technique. IET Control Theory & Applications, 11(3),390-400.

Vörös J. 2019.Identification of nonlinear block-oriented systems with backlash and saturation. Journal of Electrical Engineering, 70(2),138-144.

Li K. Peng J., Bai E. 2009.Two-Stage Mixed Discrete–Continuous Identification of Radial Basis Function (RBF) Neural Models for Nonlinear Systems. IEEE Transactions on Circuits and Systems I: Regular Papers, 56(3),630-643.

Wahlberg B., Ljung L.2018. Algorithms and performance analysis for stochastic Wiener system identification. IEEE Control Systems Letters, 2(3),471-476.

Li M., Liu X. 2018.Auxiliary Model based Least Squares Iterative Algorithms for Parameter Estimation of Bilinear Systems Using Interval-varying Measurements. IEEE Access, 2018,1-1.

Liu Q., Ding F. 2019.Auxiliary model-based recursive generalized least squares algorithm for multivariate output-error autoregressive systems using the data filtering. Circuits, Systems, and Signal Processing, 38(2),590-610.

Li Linwei is currently an instructor at School of Electrical and Information Engineering, Zhengzhou University of Light Industry, Zhengzhou, China. He received his M.S. degrees from School of Electrical Engineering and Automation, Henan Polytechnic University, Henan, China in 2015. He received the Ph.D. degree in the School of Automation, Beijing Institute of Technology, Beijing, China in 2019. His research interests include the identification and control of nonlinear system, modeling and control of servo system, adaptive control.

Xianglong Liu received the B.Eng. degree in automation and M.Eng. degree in control theory and control engineering from Henan Polytechnic University, Henan, China, and the Ph.D. degree from Beijing Jiaotong University. He is currently a lecturer with Zhengzhou University of Light Industry. His current research interests include electromagnetic tomography, precise measurement, signal processing and nondestructive testing.