Parameter Identification of Nonlinear Systems with Time-delay Based on the Multi-innovation Stochastic Gradient Algorithm

Chunming Xu

School of Mathematics and Statistics, Yancheng Teachers University, 224002, Yancheng, PR China

Copyright ©2019 by authors, all rights reserved. Authors agree that this article remains permanently open access under the terms of the Creative Commons Attribution License 4.0 International License

Abstract This paper considers the parameter identification problem of block-oriented Hammerstein nonlinear systems with time-delay. Firstly, we adopt the data filtering technique to transform the identification model so that all the parameters will be separated in the resulting identification model which has no redundant parameters. Secondly, a multi-innovation stochastic gradient algorithm is used to estimate the system parameters. The proposed method has high computational efficiency and good accuracy. Simulation results are presented to demonstrate the effectiveness of the proposed algorithm.

Keywords Parameter Identification, Stochastic Gradient Algorithm, Multi-innovation Stochastic Gradient Algorithm, Nonlinear Systems, Time-delay

oriented multi-input single-output Hammerstein nonlinear systems which have the following form:

$$A(z)y(t) = B(z)u_f(t) + v(t) \quad (1)$$

where $u(t)$ and $y(t)$ are the system input and output, and $v(t)$ is a white noise with zero mean, $u_f(t)$ is a nonlinear function, $A(z)$ and $B(z)$ are the polynomials in the unit backward shift operator $z^{-1}y(t) = y(t-1)$ which contains the parameter vectors to be identified.

Many studies have been done on the identification of nonlinear systems which include the products of the original system parameters such as (1). For example, Wang et al. presented an extended stochastic gradient identification algorithm for Hammerstein-Wiener system with colored noises and gave two methods of separating the parameters of the system[14]. Xiao et al. presented a filtering based recursive least squares identification algorithm for a input nonlinear dynamical adjustment models with memoryless nonlinear blocks followed by linear dynamical blocks[15]. However, the over-parameterized methods have many redundant estimates and the dimensions of parameter vectors will be larger. As a result, the accuracy and speed of the parameters evaluation will be affected.

Due to that time delay exists in many nonlinear systems, identification and construction of delay system models from experimental data have great value in control area. In this paper, on the basis of the work in[16], a multi-innovation stochastic gradient algorithm is proposed to identify the parameters of nonlinear systems with nonlinear systems which include the products of the original system parameters and time-delay.

The rest of this paper is organized as follows. The problem formulation and identification model are performed in section 2. Section 3 derives the parameters identification process of the nonlinear systems with time-delay. Numerical simulations are performed in section 4 to verify the effectiveness of the presented schemes, and concluding remarks are made in the final section.

1. Instruction

Parameter identification means estimating the parameters of partially unknown systems based on noisy observations, which is the foundation of many issues such as signal processing, system identification and system control[1–5]. In the last few decades, the system identification theory has been well studied and many system identification methods e.g., the instrumental variable methods[6, 7], the bias compensation methods[8, 9], the least squares methods[10, 11], the multi-innovation identification methods[12, 13] have been developed to model the dynamical systems.

In reality, many exist systems exhibit nonlinear phenomena and nonlinear system identification is currently an active research topic in parameter identification. Typical nonlinear systems include the Hammerstein system, the Wiener systems and their combinations, etc. The Hammerstein system consists of a nonlinear block plus a linear block, while the Wiener systems consists of a linear dynamic system block followed by a static nonlinearity block. The above systems have been well studied and there are a lot of parameter identification algorithms about them have been proposed. In practice, there is a class of block-

2. System description and identification model

Consider the following nonlinear system with time-delay:

$$A(z)y(t) = B(z)u_f(t - \tau) + v(t) \quad (2)$$

where $u_f(t)$ is a nonlinear function which can be approximable by a polynomial: $u_f(t) = \sum_{j=1}^m \beta_j u^j(t)$, τ is the time delay, $v(t)$ is a random white noise with zero mean and variance δ^2 , $y(t)$ is the measured system output, and $A(z)$ and $B(z)$ are the polynomials in the unit backward shift operator z^{-1} $y(t) = y(t - 1)$ with:

$$A(z) := 1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_{n_a} z^{-n_a} \quad (3)$$

$$B(z) := 1 + b_1 z^{-1} + b_2 z^{-2} + \dots + b_{n_b} z^{-n_b} \quad (4)$$

Assume that the order n_a and n_b are known and $u(t) = 0$, $v(t) = 0$ and $y(t) = 0$ as $t \leq 0$. Multiplying both sides of the equation (2) by $\frac{1}{B(z)}$ yields

$$\frac{A(z)}{B(z)}y(t) = \sum_{j=1}^m \beta_j u^j(t - \tau) + \frac{1}{B(z)}v(t) \quad (5)$$

Define the filtered output and the filtered noise as

$$y_f(t) = \frac{1}{B(z)}y(t) \quad (6)$$

$$v_f(t) = \frac{1}{B(z)}v(t) \quad (7)$$

Then we can have

$$y_f(t) = -\sum_{j=1}^{n_b} b_j y_f(t - j) + y(t) \quad (8)$$

$$v_f(t) = -\sum_{j=1}^{n_b} b_j v_f(t - j) + v(t) \quad (9)$$

On the other hand, Eq.(5) can be rewritten as

$$A(z)y_f(t) = \sum_{j=1}^m \beta_j u^j(t - \tau) + v_f(t) \quad (10)$$

Then we can get the following expression

$$y_f(t) = -\sum_{j=1}^{n_a} a_j y_f(t - j) + \sum_{j=1}^m \beta_j u^j(t - \tau) + v_f(t) \quad (11)$$

Substituting (8) and (9) into (11) gives

$$\begin{aligned} y(t) &= \sum_{j=1}^{n_b} b_j y_f(t - j) - \sum_{j=1}^{n_a} a_j y_f(t - j) \\ &+ \sum_{j=1}^m \beta_j u^j(t - \tau) - \sum_{j=1}^{n_b} b_j v_f(t - j) + v(t) \\ &= -\sum_{j=1}^{n_a} a_j y_f(t - j) + \sum_{j=1}^{n_b} b_j [y_f(t - j) - v_f(t - j)] \\ &+ \sum_{j=1}^m \beta_j u^j(t - \tau) + v(t) \end{aligned} \quad (12)$$

Define the parameter vectors

$$\theta = [a^T, b^T, d^T, \beta^T]^T \in R^n \quad (13)$$

$$a = [a_1, a_2, \dots, a_{n_a}]^T \in R^{n_a} \quad (14)$$

$$b = [b_1, b_2, \dots, b_{n_b}]^T \in R^{n_b} \quad (15)$$

$$d = [d_1, d_2, \dots, d_{n_d}]^T \in R^{n_d} \quad (16)$$

$$\beta = [\beta_1, \beta_2, \dots, \beta_{n_\beta}]^T \in R^{n_\beta} \quad (17)$$

$$p(t) = y_f(t) - v_f(t) \quad (18)$$

and the information vectors

$$\Psi_f(t) = [-y_f(t - 1), -y_f(t - 2), \dots, -y_f(t - n_a), p(t - 1), p(t - 2), \dots, p(t - n_b), u(t - \tau), u^2(t - \tau), \dots, u^m(t - \tau)] \quad (19)$$

we have

$$y(t) = \Psi_f^T(t)\theta + v(t) \quad (20)$$

Eq.(20) is the identification model.

3. The stochastic gradient (SG) algorithm

In the following, we derive a stochastic gradient algorithm for the nonlinear systems with time-delay to identify the parameter vector θ . Define the cost functions:

$$J(\theta) = (y(t) - \Psi_f^T(t)\theta)^2 \quad (21)$$

Let $\hat{\theta}(t)$ be the estimate of θ at time t . Using the gradient search and minimizing $J(\theta)$, we can obtain the following stochastic gradient algorithm of computing the estimate $\hat{\theta}(t)$:

$$\hat{\theta}(t) = \hat{\theta}(t - 1) + \frac{\Psi_f(t)}{r(t)} [y(t) - \Psi_f^T(t)\hat{\theta}(t - 1)] \quad (22)$$

$$r(t) = r(t - 1) + \lambda(t) \|\Psi_f(t)\|^2, r(0) = 1 \quad (23)$$

However, the information vector $\Psi(t)$ contains unknown inner variables $y_f(t - i)$, $i = 1, 2, \dots, n_c$, $v_f(t - i)$, $i = 1, 2, \dots, n_d$ and $v(t - i)$, $i = 1, 2, \dots, n_d$, the stochastic gradient algorithm cannot be applied directly. By means of the auxiliary model identification idea, we can replace the unknown variables by their corresponding estimates $\hat{y}_f(t - i)$, $i = 1, 2, \dots, n_a$, $\hat{v}_f(t - i)$, $i = 1, 2, \dots, n_b$ and $\hat{v}(t - i)$, $i = 1, 2, \dots, n_b$. Define

$$\hat{\Psi}_f(t) = [-y_f(t - 1), -y_f(t - 2), \dots, -y_f(t - n_a), p(t - 1), p(t - 2), \dots, p(t - n_b), u(t - \tau), u^2(t - \tau), \dots, u^m(t - \tau)] \quad (24)$$

where

$$\hat{y}_f(t) = -\sum_{j=1}^{n_b} b_j \hat{y}_f(t - j) + y(t) \quad (25)$$

$$\hat{v}_f(t) = -\sum_{j=1}^{n_b} b_j \hat{v}_f(t - j) + \hat{v}(t) \quad (26)$$

$$\hat{v}(t) = y(t) - \Psi_f^T(t)\hat{\theta}(t) \quad (27)$$

$$\hat{p}(t) = \hat{y}_f(t) - \hat{v}_f(t) \quad (28)$$

Thus the stochastic gradient algorithm can be described as follows:

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \frac{\hat{\Psi}_f(t)}{r(t)} [y(t) - \hat{\Psi}_f^T(t) \hat{\theta}(t-1)] \quad (29)$$

$$r(t) = r(t-1) + \lambda(t) \left\| \hat{\Psi}_f(t) \right\|^2, r(0) = 1 \quad (30)$$

$$\hat{y}_f(t) = - \sum_{j=1}^{n_b} b_j \hat{y}_f(t-j) + y(t) \quad (31)$$

$$\hat{v}_f(t) = - \sum_{j=1}^{n_b} b_j \hat{v}_f(t-j) + \hat{v}(t) \quad (32)$$

$$\hat{v}(t) = y(t) - \hat{\Psi}_f^T(t) \hat{\theta}(t) \quad (33)$$

$$\hat{p}(t) = \hat{y}_f(t) - \hat{v}_f(t) \quad (34)$$

The steps of computing the parameter estimate $\theta(t)$ as t increases using the SG algorithm is summarized as follows:

1. Let $t = 1$, set the initial values $\hat{\theta}(0) = \mathbf{1}_n/p_0$, where p_0 is a large number (e.g., $p_0 = 10^6$).
2. Collect the input data $u(t)$ and the output data $y(t)$.
3. Form $\hat{v}(t)$, $\hat{y}_f(t)$ and $\hat{v}_f(t)$ using (33), (31) and (32), respectively. Form $\hat{p}_f(t)$ using Eq.(34).
4. Form $\hat{\Psi}_f^T(t)$ using Eq.(24).
5. Update the parameter estimate $\hat{\theta}(t)$ using Eq.(29).
6. Increase t by 1 and goto step 3.

4. The multi-innovation stochastic gradient (MI-SG) algorithm

Let p be the data length $p \gg N_0$, and define the stacked vectors or matrix

$$\mathbf{Y}(t) = \begin{bmatrix} y(t) \\ y(t-1) \\ \vdots \\ y(t-p+1) \end{bmatrix} \quad (35)$$

$$\Phi(t) = [\hat{\Psi}_f(t), \hat{\Psi}_f(t-1), \dots, \hat{\Psi}_f(t-p+1)] \quad (36)$$

$$\mathbf{V}(t) = \begin{bmatrix} v(t) \\ v(t-1) \\ \vdots \\ v(t-p+1) \end{bmatrix} \quad (37)$$

From (4) and (5), we have

$$\mathbf{Y}(t) = \Phi^T(t) \theta + \mathbf{V}(t) \quad (38)$$

Form a cost function

$$J(\theta) = \left\| \mathbf{Y}(t) - \Phi^T(t) \theta \right\|^2 \quad (39)$$

where the norm of the matrix X is defined as $\|X\|^2 = \text{tr}[XX^T]$. Based on the multi-innovation identification theory, the multi-innovation stochastic gradient algorithm can be described as follows:

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \frac{\Phi(t)}{r(t)} [\mathbf{Y}(t) - \Phi^T(t) \hat{\theta}(t-1)] \quad (40)$$

$$r(t) = r(t-1) + \|\Phi(t)\|^2, r(0) = 1 \quad (41)$$

$$\mathbf{Y}(t) = \begin{bmatrix} y(t) \\ y(t-1) \\ \vdots \\ y(t-p+1) \end{bmatrix} \quad (42)$$

$$\Phi(t) = [\hat{\Psi}_f(t), \hat{\Psi}_f(t-1), \dots, \hat{\Psi}_f(t-p+1)] \quad (43)$$

The steps involved in computing the parameter estimate θ as t increases using the MI-SG algorithm is summarized as follows:

1. Let $t = 1$, set the initial values $\hat{\theta}(0) = \mathbf{1}_n/p_0$, where p_0 is a large number (e.g., $p_0 = 10^6$).
2. Collect the input data $u(t)$ and the output data $y(t)$.
3. Form $\hat{v}(t)$, $\hat{y}_f(t)$ and $\hat{v}_f(t)$ using (33), (31) and (32), respectively. Form $\hat{p}_f(t)$ using Eq.(34).
4. Form $\hat{\Psi}_f^T(t)$ using Eq.(24).
5. Form $\mathbf{Y}(t)$ and $\Phi(t)$ using (43) and (44), respectively.
6. Update the parameter estimate $\hat{\theta}(t)$ using Eq.(40).
7. Increase t by 1 and goto step 3.

5. Example

This section provides an example to show the effectiveness of the MI-SG algorithm for the nonlinear systems with time-delay, compared with the SG algorithm. Consider the following nonlinear system:

$$A(z)y(t) = B(z)u_f(t-\tau) + v(t) \quad (44)$$

$$A(z) := 1 - 0.55z^{-1} + 0.68z^{-2} \quad (45)$$

$$B(z) := 0.80z^{-1} + 0.31z^{-2} \quad (46)$$

$$u_f(t) = \beta_1 u(t) + \beta_u^2(t) = -0.15u(t) + 0.22u^2(t) \quad (47)$$

Then the parameter vector θ can be expressed below:

$$\theta = [-0.55, 0.68, 0.8, 0.31, -0.15, 0.22]^T \quad (48)$$

In simulation, the inputs $u(t)$ is taken as an uncorrelated persistent excitation signal sequence with zero mean and unit variance, and $v(t)$ as a white noise sequence with zero mean and variance $\delta_v^2 = 0.2^2$. We apply the SG algorithm and MI-SG algorithm ($p=4$ and 8) to estimate the parameters of this system.

Numerical results are displayed in tables 1, 2, 3 and figure 1. The parameter estimates and their errors with different innovation length are shown in tables 1, 2 and 3, and the parameter estimation errors $\delta = \|\hat{\theta}(t) - \theta\|/\|\theta\|$ vs. t are shown in figure 1. In addition, figure 2 shows the time evolution curves of the system parameters vs. t of the MI-SG method with $p=8$.

Table 1. The SG estimates and errors.

t	a_1	a_2	b_1	b_2	β_1	β_2	$\delta(\%)$
100	-0.14310	0.11216	0.47317	0.22177	0.04850	0.02966	65.68690
200	-0.15713	0.13618	0.51076	0.23008	0.01729	0.03477	61.90697
500	-0.19066	0.19321	0.56333	0.24181	-0.02461	0.05657	54.63387
1000	-0.19657	0.21101	0.59483	0.24984	-0.03919	0.06592	52.09505
2000	-0.21258	0.22901	0.62456	0.25523	-0.05101	0.07096	49.36333
3000	-0.22247	0.25131	0.63931	0.26191	-0.06227	0.08225	46.90712
True values	-0.55000	0.68000	0.80000	0.31000	-0.15000	0.22000	0.00000

Table 2. The MI-SG(p=4) estimates and errors.

t	a_1	a_2	b_1	b_2	β_1	β_2	$\delta(\%)$
100	-0.32968	0.41749	0.76298	0.29920	-0.00427	0.03570	33.28421
200	-0.37303	0.46399	0.78453	0.30100	-0.07585	0.05127	26.72472
500	-0.45152	0.54888	0.79611	0.31735	-0.12354	0.11175	15.82760
1000	-0.45498	0.56107	0.80655	0.31971	-0.14123	0.12602	14.31642
2000	-0.48609	0.56367	0.80725	0.29665	-0.14648	0.14769	12.12000
3000	-0.48873	0.59291	0.80472	0.30797	-0.14999	0.16727	9.48666
True values	-0.55000	0.68000	0.80000	0.31000	-0.15000	0.22000	0.00000

Table 3. The MI-SG(p=8) estimates and errors.

t	a_1	a_2	b_1	b_2	β_1	β_2	$\delta(\%)$
100	-0.40943	0.61904	0.81142	0.28412	-0.04854	0.02934	21.24124
200	-0.50181	0.62366	0.80782	0.29376	-0.13550	0.05995	14.18953
500	-0.57338	0.68181	0.80418	0.32645	-0.15458	0.15748	5.50753
1000	-0.55496	0.67041	0.81802	0.32351	-0.17080	0.16283	5.24515
2000	-0.57938	0.64729	0.81008	0.28240	-0.16838	0.18978	5.07445
3000	-0.55946	0.67635	0.80468	0.30806	-0.16271	0.21254	1.48267
True values	-0.55000	0.68000	0.80000	0.31000	-0.15000	0.22000	0.00000

Figure 1. The parameter estimation errors δ vs. t .

Figure 2. The time evolution curves of the system parameters vs. t .

The experiments has been conducted to demonstrate the properties of the proposed algorithm. Some attractive points can be summarized as follows:

- (1) The parameter estimation errors of the three algorithms become smaller and smaller as t increasing.
- (2) MI-SG performers better than SG because it can consider more data information.
- (3) When the innovation length p becomes longer, the better identification results will be obtained.

6. Conclusions

This paper investigates the parameter identification of the nonlinear system with time-delay using the data filtering technique and the multi-innovation stochastic gradient algorithm. Based on the data filtering technique, the original identification model is converted into a new model which has no cross-products of the parameters and is easier to be identified. The simulation results indicate that the proposed method is effective. The presented method can also be used to solve other relevant system identification problems.

Acknowledgements

The authors are grateful to the anonymous reviewers for their valuable comments and suggestions that have helped to improve the presentation of the paper. This work is partially supported by the National Natural Science Foundation of China (Grant No.11771376).

REFERENCES

- [1] J.P. Norton. An Introduction to Identification, Academic Press, London and New York, 1986.
- [2] T. Soderstrom, P. Stoica. System identification, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1989.
- [3] L. Ljung. System Identification. Theory for the User, second ed., Prentice-Hall, Inc., Englewood Cliffs, NJ, 1999.
- [4] R.V. Herpen, O. Bosgra, T. Oomen. Bi-Orthonormal Polynomial Basis Function Framework With Applications in System Identification, IEEE Transactions on Automatic Control, vol.61, no.11, 3285-3300, 2012.
- [5] G. Pillonetto, T. Chen, A. Chiuso, G.D. Nicolao, L. Ljung. Regularized linear system identification using atomic, nuclear and kernel-based norms, Automatica, vol.69, no.3, 137-149, 2016.
- [6] T. Söderström, P. Stoica. Instrumental variable methods for system identification, Circuits Systems and Signal Processing, vol.21, no.1, 1-9, 2002.
- [7] Walid Allafi, Kotub Uddin, Cheng Zhang, Raja Mazuir Raja Ahsan Sha, James Marco. On-line scheme for parameter estimation of nonlinear lithium ion battery equivalent circuit models using the simplified refined instrumental variable method for a modified Wiener continuous-time model. Aiche Journal, vol.204, no.15, 497-508, 2017.
- [8] M.J. Sang, P.G. Park. Stabilization of a Bias-Compensated Normalized Least-Mean-Square Algorithm for Noisy Inputs. IEEE Transactions on Signal Processing, vol.65, no.11, 2949-2961, 2017.
- [9] Z. Zheng, Z. Liu, L. Lu. Bias-compensated robust set-membership NLMS algorithm against impulsive noises and noisy inputs. Electronics Letters, vol.53, no.16, 1100-1102, 2017.
- [10] H. Hu, R. Ding. Least squares based iterative identification algorithms for input nonlinear controlled autoregressive systems based on the auxiliary model. Nonlinear Dynamics, vol.76, no.1, 777-784, 2013.
- [11] F. Ding, Y. Wang, J. Ding. Recursive least squares parameter identification algorithms for systems with colored noise using the filtering technique and the auxiliary model. Digital Signal Processing, vol.37, no.1, 100-108, 2015.
- [12] S. Cheng, Y. Wei, D. Sheng, Y. Chen, Y. Wang. Identification for Hammerstein nonlinear ARMAX systems based on multi-innovation fractional order stochastic gradient, Signal Processing, vol.142, no.1, 1-10, 2017.
- [13] F. Ding, H. Chen, M. Li. Multi-innovation least squares identification methods based on the auxiliary model for MISO systems. Applied Mathematics and Computation, vol.187, no.2, 658-668, 2007.
- [14] D.Q. Wang, F. Ding. Extended stochastic gradient identification algorithms for Hammerstein-Wiener ARMAX systems, Computers and Mathematics with Applications. vol.56, no.12, 3157-3164, 2008.
- [15] Y.S. Xiao, N. Yue. Parameter estimation for nonlinear dynamical adjustment models. Mathematical and Computer Modelling, vol.54, no.5, 1561C1568, 2011.
- [16] Y.W. Mao, F. Ding. A novel parameter separation based identification algorithm for Hammerstein systems. Applied Mathematics Letters, vol.60, no.10, 21-27, 2016.