Identification of Hammerstein Systems with a Recursive Algorithm

Li, Zhenqiang
Department of Electrical and Electronic Systems Engineering, Graduate School of Information Science and Electrical Engineering, Kyushu University : Graduate Student

Kanae, Shunshoku
Department of Electrical and Electronic Systems Engineering, Faculty of Information Science and Electrical Engineering, Kyushu University

Yang, Zi-Jiang
Department of Electrical and Electronic Systems Engineering, Faculty of Information Science and Electrical Engineering, Kyushu University

Wada, Kiyoshi
Department of Electrical and Electronic Systems Engineering, Faculty of Information Science and Electrical Engineering, Kyushu University

https://doi.org/10.15017/1517946

出版情報：九州大学大学院システム情報科学紀要. 14 (1), pp.1-5, 2009-03-26. 九州大学大学院システム情報科学研究院
バージョン：
権利関係：

九州大学大学院
システム情報科学紀要
第14巻　第1号　平成21年3月

Research Reports on Information Science and
Electrical Engineering of Kyushu University
Vol.14, No.1, March 2009

Identification of Hammerstein Systems with a Recursive Algorithm

Zhenqiang LI*, Shunshoku KANAE**, Zi-Jiang YANG** and Kiyoshi WADA**

(Received December 8, 2008)

Abstract: This paper deals with the problem of estimating the parameters of Hammerstein systems based on recursive least squares method. Hammerstein systems can be considered as one kind of nonlinear systems, it is applied in many fields. Several identification methods of Hammerstein systems have been developed. We utilize the recursive algorithm to identify the system and the obtained simulation results show the effectiveness of this approach.

Keywords: Hammerstein systems, Nonlinear system identification, ARX model, Parameter identification, Recursive least squares method

1. Introduction

Identification of linear system has been a mature technique. However, the methods are very difficult to identify nonlinear system, because there are many high order items, cross items and other nonlinear function relationship between inputs and outputs. In fact, most sensors and actuators have multiple nonlinear features, the research on the identification of nonlinear system is very attractive. Without any priori structural information, nonlinear system identification is an intractable task. There are some usual ways to identify nonlinear system, such as artificial neural network, functional series like as Volterra series and Wiener series, block models. Artificial neural network can approximate any nonlinear characteristics, but many neurons and layers will be applied, and it spends long time to learn and train. Functional series such as Volterra series and Wiener series can describe any nonlinear dynamic systems. The kernel functions of these series are a series of infinite sum of multiple convolution integrals, however, estimating all coefficients of the kernel functions is unfeasible. In fact, only coefficients of finite order kernel functions can be identified. Usually, the structure of nonlinear system can be specified as a certain type, such as Hammerstein model, Wiener model and their combinations. These kinds of nonlinear models are called block-oriented models.

Hammerstein systems are the most widely applied nonlinear dynamic models. It assumes that a nonlinear system is divided into two parts: the static nonlinear part and the dynamic linear part. The Hammerstein systems consist of a static memoryless nonlinear block followed by a linear dynamic block; Wiener systems have a linear dynamic block followed by a static memoryless nonlinear block. Many researchers have proposed lots of methods on the parametric identification of Hammerstein systems, these methods include: the iterative method[1)] [6)], the over-parameterization method[2)], the blind method[3)], the frequency domain method[4)], etc.

The iterative method is the basic approach of parameter estimation. All the parameters are divided into two subset, the optimal values are found in the first set while the second set is fixed. Then the first set is fixed to find the optimal values in the second set. The estimate results show that it is a good algorithm, but its convergence is a problem. The over-parameterization method is often applied to estimate linear system. For the Hammerstein systems that the unknown nonlinear block is parameterized linearly with unknown parameters. The order of the new linear system will be so large. The blind method is based on using the output measurements, once the linear part is obtained by fast sampling at the output, identification of the nonlinear part can be carried out in a number of ways[3)]. By means of Fourier transform, the frequency domain method makes the nonlinear part expand to a series of Fourier function, then the linear part and the nonlinear part can be identified. The problem of convergence rate of Fourier series is a topic[4)].

The key problem of the Hammerstein systems identification is how to describe the nonlinear block. For a static memoryless nonlinear model, if it can be represented by a function based on a certain series of basic function, for example, Fourier function, spline function, wavelet function, or polynomial function,

* Department of Electrical and Electronic Systems Engineering, Graduate Student
** Department of Electrical and Electronic Systems Engineering

etc. According to the described function which is parameterized in line by unknown parameters, the coefficients of the function need to be estimated by some identification algorithm. In this paper, we discuss that parameter identification of Hammerstein systems with a recursive method in open loop.

This paper is organized as follows: Section 2 introduces the Hammerstein model and problem formulation. A recursive identification algorithm is derived in Section 3. Section 4 provides an illustrative example to prove the effectiveness of the algorithm. Finally, some conclusions are obtained in Section 5.

2. Problem Formulation

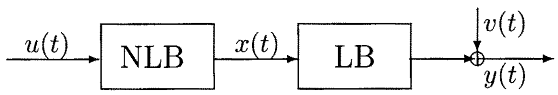

Fig. 1 Hammerstein systems.

Consider the Hammerstein systems shown as **Fig.1**. The nonlinear block of the systems can be represented a polynomial with a known order in the input signal as follows

$$x(t) = p_1 u(t) + p_2 u^2(t) + ... + p_i u^i(t). \tag{1}$$

The linear dynamic block can be described as

$$A(q^{-1})y(t) = B(q^{-1})x(t) + v(t). \tag{2}$$

Where $u(t)$, $y(t)$ and $v(t)$ are input, output and noise, respectively. $x(t)$ denotes an immeasurable internal signal and $x(t) = f(u)$, which is the output of the nonlinear block and just as the input of the linear block $G(q^{-1})$. The q^{-1} is the pure time delay of the system $[q^{-1}y(t) = y(t-1)]$, $A(q^{-1})$ and $B(q^{-1})$ are scalar polynomials in the shift operator q^{-1}:

$$\begin{aligned} A(q^{-1}) &= 1 + a_1 q^{-1} + a_2 q^{-2} + ... + a_n q^{-n}, \\ B(q^{-1}) &= 1 + b_1 q^{-1} + b_2 q^{-2} + ... + b_n q^{-n}. \end{aligned}$$

Parameter identification is to estimate all coefficients of the nonlinear block $f(u)$ and linear block $G(q^{-1})$ by measuring inputs and outputs of the system. For some nonzero and finite constant k, any pair $(kf(u), G(q^{-1})/k)$ will produce the same input and output measurements. In other words, any identification scheme cannot distinguish between $(f(u), G(q^{-1}))$ and $(kf(u), G(q^{-1})/k)$ [3]. Therefore, in order to obtain an unique estimate, without loss of generality, either gain of $f(u)$ or $G(q^{-1})$ must be fixed[3]. Where the first coefficient p_1 of function $f(u)$ is defined as one. The Hammerstein model can be described as follows

$$A(q^{-1})y(t) = B(q^{-1})f(u(t)) + v(t). \tag{3}$$

The output equation shows in the form:

$$y(t) = -\sum_{i=1}^{n} a_i y(t-i) + \sum_{i=1}^{n}\sum_{j=1}^{m} b_i p_j u^j(t-i) + v(t).$$

By sampling the available input and output data $u(t), y(t)$, the parameters of the Hammerstein systems a_i, b_i, p_j, can be estimated with the following algorithm.

3. Identification Algorithm

Define the following parameter vector and data vector

$$\Theta = \begin{bmatrix} \mathbf{a} \\ p_1\mathbf{b} \\ p_2\mathbf{b} \\ \vdots \\ p_m\mathbf{b} \end{bmatrix} \in R^r, \Psi(t) = \begin{bmatrix} \mathbf{Y}(t) \\ \mathbf{U}_1(t) \\ \mathbf{U}_2(t) \\ \vdots \\ \mathbf{U}_m(t) \end{bmatrix} \in R^r, \tag{4}$$

$r := (m+1)n$,

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \in R^n, \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \in R^n, \tag{5}$$

$$\mathbf{Y}(t) = \begin{bmatrix} -y(t-1) \\ -y(t-2) \\ \vdots \\ -y(t-n) \end{bmatrix} \in R^n, \ \mathbf{Y}_N = \begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(N) \end{bmatrix}, \tag{6}$$

$$\mathbf{U}_j(t) = \begin{bmatrix} U^j(t-1) \\ U^j(t-2) \\ \vdots \\ U^j(t-n) \end{bmatrix} \in R^n, \ \Phi_N = \begin{bmatrix} \Psi(1) \\ \Psi(2) \\ \vdots \\ \Psi(N) \end{bmatrix} \tag{7}$$

$j = 1, 2, ..., m$. Then we have

$$y(t) = \Psi^T(t)\Theta + v(t). \tag{8}$$

Since $v(t)$ is a white noise with zero mean and finite variance, we define $\hat{\Theta}$ as the estimated value of Θ, the output prediction value can be described as follows

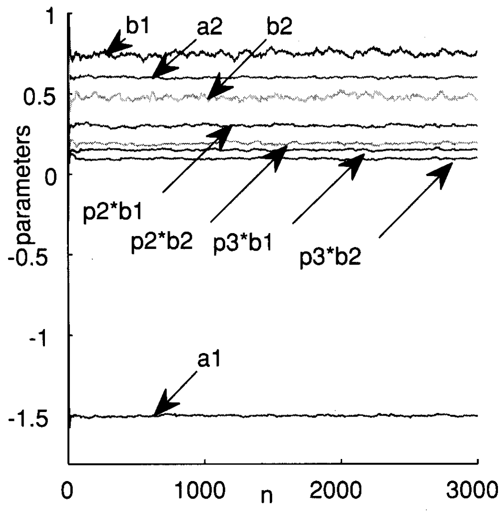

Fig. 2 Parameter estimates ($NSR = 10\%$, $\lambda = 0.98$).

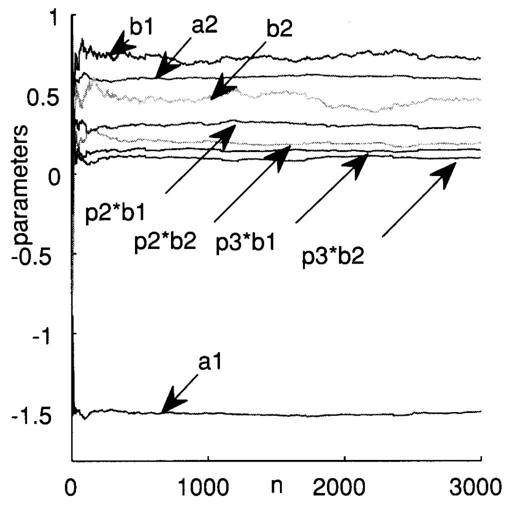

Fig. 3 Parameter estimates ($NSR = 50\%$, $\lambda = 0.998$).

$$\hat{y}(t) = \Psi^T(t)\hat{\Theta}. \tag{9}$$

The quadratic output prediction error criterion

$$\mathbf{J}(\hat{\Theta}) := \sum_{i=1}^{r}[y(i) - \hat{y}(i)]^2 = \sum_{i=1}^{r}[y(i) - \Psi^T(i)\hat{\Theta}]^2.$$

In order to minimize the error function

$$\frac{\partial \mathbf{J}(\hat{\Theta})}{\partial \hat{\Theta}} = 0. \tag{10}$$

We obtain the least-squares estimation

$$\hat{\Theta}_N = [\Phi_N^T \Phi_N]^{-1}\Phi_N^T \mathbf{Y}_N, \tag{11}$$

A recursive identification algorithm of estimating θ can be derived as follows

$$\hat{\Theta}_{N+1} = \hat{\Theta}_N + \frac{Q_N \Psi(N+1)[y(N+1) - \Psi^T(N+1)\hat{\Theta}_N]}{\lambda + \Psi^T(N+1)Q_N\Psi(N+1)}$$

$$Q_{N+1} = \frac{1}{\lambda}[Q_N - \frac{Q_N\Psi(N+1)\Psi^T(N+1)Q_N}{\lambda + \Psi^T(N+1)Q_N\Psi(N+1)}],$$

$$Q_N = [\Phi_N^T\Phi_N]^{-1}, \quad Q(0) = \mu\mathbf{I}, \ 0 < \mu < \infty.$$

Where we introduce some notation, the symbol $\mathbf{I}$ stands for an identity matrix of appropriate size; $\lambda \leq 1$ is called forgetting factor; T is a definition of the matrix transpose; the $Q(0)$ is an initialization which is a positive real matrix e.g. $Q(0) = 10^6\mathbf{I}$.

According to the parameter $p_1 = 1$, the estimated parameters of $\mathbf{a}$ and $\mathbf{b}$, $\hat{\mathbf{a}} = [\hat{a}_1 \ \ \hat{a}_2 \ \ ... \ \ \hat{a}_n]^T$, $\hat{\mathbf{b}} = [\hat{b}_1 \ \ \hat{b}_2 \ \ ... \ \ \hat{b}_n]^T$ can be directly obtained from the first, second element of $\hat{\Theta}$, respectively. Let $\hat{\Theta}_j$ be the jth element of $\hat{\Theta}$, then $\hat{\Theta}_j = \widehat{p_j\mathbf{b}}$. There is a large amount of redundancy in the estimation of parameters of the nonlinear static part. By utilizing the average method, the estimates can be found

$$\hat{p}_j = \frac{1}{n}\sum_{i=1}^{n}\frac{\widehat{p_j b_i}}{\hat{b}_i} \quad j = 2, 3, ..., m. \tag{12}$$

4. Example

An example is given to demonstrate the effectiveness of the recursive algorithm. The nonlinear static block is

$$\begin{aligned} x(t) &= p_1u(t) + p_2u^2(t) + p_3u^3(t) \\ &= u(t) + 0.4u^2(t) + 0.2u^3(t). \end{aligned}$$

The linear dynamic model is an ARX model

$$A(q^{-1})y(t) = B(q^{-1})f(u(t)) + v(t),$$

$$\begin{aligned} A(q^{-1}) &= 1 + a_1q^{-1} + a_2q^{-2} \\ &= 1 - 1.50q^{-1} + 0.6q^{-2}, \end{aligned}$$

$$B(q^{-1}) = b_1q^{-1} + b_2q^{-2} = 0.75q^{-1} + 0.48q^{-2},$$

$$\Theta = [a_1 \ a_2 \ b_1 \ b_2 \ p_2 \ p_3]^T.$$

These parameters will be estimated. 20 Monte Carlo runs are calculated. For each Monte Carlo run, the input and the noise are uniformly distri-

Table 1 The parameter estimates (NSR = 10%, λ = 0.98).

Θ	a_1	a_2	b_1
true	-1.5000	0.6000	0.7500
N=100	-1.4130	0.5327	0.7142
N=200	-1.4543	0.5649	0.7308
N=500	-1.4812	0.5885	0.7421
N=1000	-1.4913	0.5932	0.7428
N=2000	-1.4981	0.5994	0.7432
N=3000	-1.4986	0.5996	0.7444
Θ	b_2	p_2	p_3
true	0.4800	0.4000	0.2000
N=100	0.4628	0.4376	0.1888
N=200	0.4762	0.4225	0.2008
N=500	0.4816	0.4096	0.2018
N=1000	0.4795	0.4085	0.1988
N=2000	0.4783	0.4058	0.2014
N=3000	0.4776	0.4042	0.2018

Table 2 The parameter estimates (NSR = 50%, λ = 0.998).

Θ	a_1	a_2	b_1
true	-1.5000	0.6000	0.7500
N=100	-1.4094	0.5311	0.6902
N=200	-1.4495	0.5621	0.7168
N=500	-1.4788	0.5839	0.7363
N=1000	-1.4912	0.5934	0.7351
N=2000	-1.5013	0.6032	0.7348
N=3000	-1.5003	0.6018	0.7371
Θ	b_2	p_2	p_3
true	0.4800	0.4000	0.2000
N=100	0.4468	0.4757	0.1476
N=200	0.4755	0.4468	0.1904
N=500	0.4858	0.4201	0.2002
N=1000	0.4796	0.4195	0.1949
N=2000	0.4767	0.4136	0.2024
N=3000	0.4748	0.4097	0.2038

bution, respectively. The input signal must be a persistent excitation signal. Since the white noise is a persistent excitation signal[5)], we select the standard white noise as input $u(t)$, the disturbance $v(t)$ is also a white noise with zero mean and finite variance σ_v^2. The $v(t)$ is independent of input $u(t)$. For the input of linear part, that is a three-order polynomial of the input $u(t)$. We apply the recursive algorithm to estimate the parameters by means of **MATLAB** in two different noise case. The estimated parameters are shown as **Fig.2** and **Fig.3**. Where NSR is defined as the ratio of noise to signal, it is the square root of the ratio of noise and output variances. Since the estimated parameters are redundant, by computing with (12), the estimated results are filled in **Table1** and **Table2** with the different data length N. From these data, we compute the relative parameter estimation errors δ_1 and δ_2 for the different data length in two noise cases. Where δ_1 and δ_2 are the errors while $NSR = 10\%$ and $NSR = 50\%$, respectively. Shown as **Table3**.

From **Figs.2-4** and **Tables1-3**, we can give the following remarks:

- Rate of convergence of the parameter estimations is slowly at a high noise case.
- The errors of the parameter estimations are getting smaller with the data length increased.
- We will select the forgetting factor which is near to one, while estimating the parameters under a high noise condition.

In the implementation of the recursive algorithm, the parameter estimate fluctuations appear in the

Table 3 The error of estimated parameter.

N	$\delta_1(\%)$	$\delta_2(\%)$	N	$\delta_1(\%)$	$\delta_2(\%)$
100	6.5097	8.5125	1000	0.8297	1.4416
200	3.4184	4.5194	2000	0.4932	1.1113
500	1.4174	1.9247	3000	0.4062	0.9184

early steps and become stable. When the output noise is with a big variance, the fluctuation phenomenon is very severe. It will result in a bad estimation of nonlinear block in Hammerstein systems. We can select different forgetting factor to obtain good parameter estimates. Finally note that there is no general proof of convergence for the recursive algorithm. The experiment provides good results of the parameter estimates. They show the effectiveness of this approach.

5. Conclusion

Since nonlinear systems have more and more types, such as dead zone, saturation etc, identification of nonlinear dynamic systems is a difficult problem. Though many identification approaches have been proposed, none universal method can be applied to estimate nonlinear systems.

In this paper, when the nonlinear part of the Hammerstein systems can be described as a known-order polynomial, we utilize the recursive method to estimate it. The simulation results indicate when the length of the sampled data is to a certain amount, the identification algorithm has good convergence property.

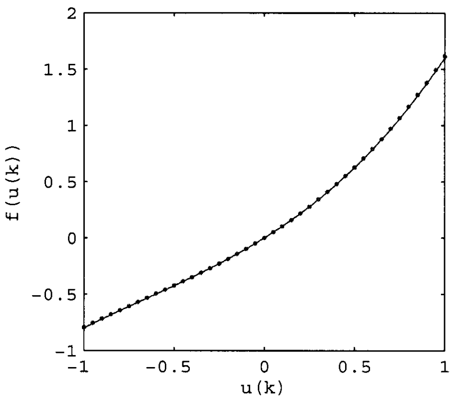

Fig. 4 The nonlinear function: actual (solid) and estimated (dotted).

References

1) K.S. Narendra and P.G. Gallman, "An iterative method for the identification of nonlinear systems using the Hammerstein model," *IEEE Trans. Autom. Control,* vol.AC-11, no.3, pp.546-550, Jul.1966.
2) E.W. Bai, "An optimal two-stage identification algorithm for Hammerstein-Wiener nonlinear systems," *Automatica*, vol.34, no.3, pp.333-338, 1998.
3) E.W. Bai, "A blind approach to Hammerstein model identification" *IEEE Trans. signal processing*, vol.50, no.7, pp.1610-1619, 2002.
4) E.W. Bai, "Frequency domain identification of Hammerstein models," *IEEE Trans. Autom. Control,* vol.48, no.4, pp.530-542, 2003.
5) L. Ljung, *System identification: theory for the users.* (2nd ed). Upper Saddle River:Prentice-Hall. 1999.
6) A.H. Falkner, "Iterative technique in the identification of a non-linear system" *International Journal of Control,* 48(1), 385-396, 1988.