Stochastic Gradient Algorithm for Hammerstein Systems with Piece-Wise Linearities*

Jing Chen[1,2] and Jia Chen[1]

[1] Wuxi Professional College of Science and Technology, Wuxi, P.R. China 214028
chenjing1981929@126.com
[2] Control Science and Engineering Research Center, Jiangnan University, Wuxi, P.R. China 214122
302774418@qq.com

Abstract. This paper studies a stochastic algorithm for Hammerstein systems with piece-wise linearities. By using a switching function, the model of the nonlinear Hammerstein systems be changed to an identification model, then based on the derived model, a stochastic gradient identification algorithm is used to estimate all the unknown parameters of the systems. An example is provided to show the effectiveness of the proposed algorithm.

Keywords: Piece-wise linearity, Stochastic gradient, Parameter estimation, Hammerstein system.

1 Introduction

Hammerstein systems consist of a static nonlinear block followed by a linear dynamic block which are widely used in many areas, e.g., nonlinear filtering, actuator saturations, audio-visual processing, signal analysis. There exists a lot of work on identification of these nonlinear systems [1–6]. Some work assumed that the nonlinearity is the polynomial nonlinearity [6–8], others assumed that the nonlinearity is the hard nonlinearity [2, 3, 9–12, 14]. Hard nonlinearity cannot be written as an analytic function of the input and is more common in engineering practice. Recently, identification of Hammerstein systems with hard nonlinearity has been received much attention [3, 9, 10, 13–15]. For example, Bai used a deterministic approach and the correlation analysis method to estimate the parameters of systems with hard input nonlinearities [9]. Chen proposed a novel estimation algorithm for dual-rate Hammerstein systems with preload nonlinearity [13], and studied identification problems for Hammerstein systems with saturation and dead-zone nonlinearities [3].

This paper deals with the identification of Hammerstein systems with piece-wise linearities. By using the switching function, the model of the Hammerstein systems can be turned into an identification model, then based on the derived

* This work was supported by the National Natural Science Foundation of China.

K. Li et al. (Eds.): ICSEE 2013, CCIS 355, pp. 241–247, 2013.

model, a stochastic gradient algorithm (SG) is proposed to estimate the unknown parameters of the systems.

Briefly, the paper is organized as follows. Section 2 describes the piece-wise linearities and derives an identification model. Section 3 studies estimation algorithms for the identification model. Section 4 provides an illustrative example. Finally, concluding remarks are given in Section 5.

2 The Piece-Wise Linearities

Consider a Hammerstein system

$$A(z)y(t) = B(z)f(u(t)) + v(t), \tag{1}$$

where $y(t)$ is the system output, $u(t)$ is the system input, and $v(t)$ is a stochastic white noise with zero mean, and $A(z)$ and $B(z)$ are polynomials in the unit backward shift operator $[z^{-1}y(t) = y(t-1)]$ and

$$\begin{aligned} A(z) &:= 1 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_n z^{-n}, \\ B(z) &:= b_1 z^{-1} + b_2 z^{-2} + b_3 z^{-3} + \cdots + b_n z^{-n}. \end{aligned}$$

The nonlinear input $f(u(t))$ is a piece-wise linearity which is shown in Figure 1 and can be expressed as

$$f(u(t)) = \begin{cases} m_1 u(t), & u(t) \geq 0, \\ m_2 u(t), & u(t) < 0, \end{cases}$$

where m_1 and m_2 are the corresponding segment slopes.

Define a switching function,

$$h(t) := h[u(t)] = \begin{cases} \frac{1}{2}, & u(t) \geq 0, \\ -\frac{1}{2}, & u(t) < 0. \end{cases}$$

Then the output $y(t)$ can be written as

$$f(u(t)) = (m_1 - m_2)u(t)h(u(t)) + \frac{1}{2}(m_1 + m_2)u(t), \tag{2}$$

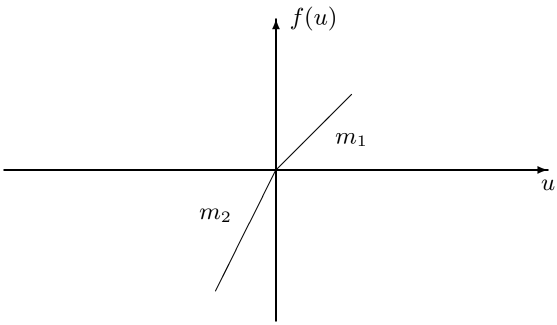

Fig. 1. The piece-wise linearity

and Equation (1) can be written as

$$\begin{aligned} A(z)y(t) = B(z)((m_1 - m_2)u(t)h(u(t)) \\ + \frac{1}{2}(m_1 + m_2)u(t)) + v(t). \end{aligned} \tag{3}$$

From (3), we can see that the output $y(t)$ of the nonlinear block can be written as an analytic function of the input.

3 The Estimation Algorithms

Define the parameter vector $\boldsymbol{\theta}$ and the information vector $\boldsymbol{\varphi}(t)$ as

$$\begin{aligned} \boldsymbol{\theta} :=& [b_1(m_1 - m_2), b_2(m_1 - m_2), b_3(m_1 - m_2), \cdots, \\ & b_n(m_1 - m_2), \frac{1}{2}b_1(m_1 + m_2), \frac{1}{2}b_2(m_1 + m_2), \\ & \frac{1}{2}b_3(m_1 + m_2), \cdots, \frac{1}{2}b_n(m_1 + m_2), \\ & a_1, a_2, a_3, \cdots, a_n]^{\mathrm{T}} \in \mathbb{R}^{3n}, \\ \boldsymbol{\varphi}(t) :=& [u(t-1)h(t-1), u(t-2)h(t-2), \\ & u(t-3)h(t-3), \cdots, u(t-n)h(t-n), \\ & u(t-1), u(t-2), u(t-3), \cdots, \\ & u(t-n), -y(t-1), -y(t-2), \cdots, \\ & -y(t-n)]^{\mathrm{T}} \in \mathbb{R}^{3n}, \end{aligned}$$

gets

$$y(t) = \boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\theta} + v(t). \tag{4}$$

If $\boldsymbol{\theta}$ has been estimated, none of the identification schemes can distinguish $b_i, i = 1, 2, 3, \cdots, n$ and $m_i, i = 1, 2$ from the estimated $\boldsymbol{\theta}$. Therefore, to get a unique parameterization, in this paper, we adopt the assumption that the first coefficient b_1 equals 1, i.e., $b_1 = 1$.

The parameter vector $\boldsymbol{\theta}$ and the information vector $\boldsymbol{\varphi}(t)$ be defined as

$$\begin{aligned} \boldsymbol{\theta} :=& [(m_1 - m_2), b_2(m_1 - m_2), b_3(m_1 - m_2), \\ & \cdots, b_n(m_1 - m_2), \frac{1}{2}(m_1 + m_2), \\ & \frac{1}{2}b_2(m_1 + m_2), \frac{1}{2}b_3(m_1 + m_2), \\ & \cdots, \frac{1}{2}b_n(m_1 + m_2), a_1, \\ & a_2, a_3, \cdots, a_n]^{\mathrm{T}} \in \mathbb{R}^{3n}, \qquad (5) \\ \boldsymbol{\varphi}(t) :=& [u(t-1)h(t-1), u(t-2)h(t-2), \\ & u(t-3)h(t-3), \cdots, u(t-n)h(t-n), \end{aligned}$$

$$u(t-1), u(t-2), u(t-3), \cdots, u(t-n),$$
$$-y(t-1), -y(t-2), \cdots,$$
$$-y(t-n)]^{\mathrm{T}} \in \mathbb{R}^{3n}, \quad (6)$$

Using the following SG algorithm to estimate the parameter vector $\boldsymbol{\theta}$ in (5):

$$\hat{\boldsymbol{\theta}}(t) = \hat{\boldsymbol{\theta}}(t-1) + \frac{\boldsymbol{\varphi}(t)}{r(t)}(y(t) - \boldsymbol{\varphi}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}(t-1)), \quad (7)$$

$$\begin{aligned} \boldsymbol{\varphi}(t) = [&u(t-1)h(t-1), u(t-2)h(t-2), \\ &u(t-3)h(t-3), \cdots, u(t-n)h(t-n) \\ &, u(t-1), u(t-2), u(t-3), \cdots, u(t-n), \\ &-y(t-1), -y(t-2), \cdots, -y(t-n)]^{\mathrm{T}}, \end{aligned} \quad (8)$$

$$r(t) = r(t-1) + \|\boldsymbol{\varphi}(t)\|^2,\ r(0) = 1. \quad (9)$$

where $\frac{1}{r(t)}$ is the step-size and the norm of matrix X is defined by $\|X\|^2 := \mathrm{tr}[XX^{\mathrm{T}}]$.

The convergence of the SG algorithm is relatively slower compared with the recursive least squares algorithm. In order to improve the tracking performance of the SG algorithm, we can introduce a λ in the SG algorithm to get the SG algorithm with a forgetting factor (the FF-SG algorithm for short) as follows:

$$\hat{\boldsymbol{\theta}}(t) = \hat{\boldsymbol{\theta}}(t-1) + \frac{\boldsymbol{\varphi}(t)}{r(t)}(y(t) - \boldsymbol{\varphi}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}(t-1)), \quad (10)$$

$$\begin{aligned} \boldsymbol{\varphi}(t) = [&u(t-1)h(t-1), u(t-2)h(t-2), \\ &u(t-3)h(t-3), \cdots, u(t-n)h(t-n), \\ &u(t-1), u(t-2), u(t-3), \cdots, u(t-n), \\ &-y(t-1), -y(t-2), \cdots, -y(t-n)]^{\mathrm{T}} \end{aligned} \quad (11)$$

$$\begin{aligned} r(t) = &\lambda r(t-1) + \|\boldsymbol{\varphi}(t)\|^2, \\ &0 < \lambda < 1,\ r(0) = 1. \end{aligned} \quad (12)$$

4 Example

Consider the following linear dynamic block,

$$[1 - 0.1q^{-1}]y(t) = [q^{-1} + 1.2q^{-2}]f(u(t)) + v(t),$$

the input $\{u(t)\}$ is taken as a persistent excitation signal sequence with zero mean and unit variance, and $\{v(t)\}$ is taken as a white noise sequence with zero mean and variance $\sigma^2 = 0.10^2$, the piece-wise linearity is shown in Figure 1 and with parameters: $m_1 = 1$, $m_2 = 0.8$. Then we have

$$\begin{aligned} \boldsymbol{\theta} = [&m_1 - m_2, b_2(m_1 - m_2), 0.5(m_1 + m_2), \\ &0.5b_2(m_1 + m_2), a_1]^{\mathrm{T}} \end{aligned}$$

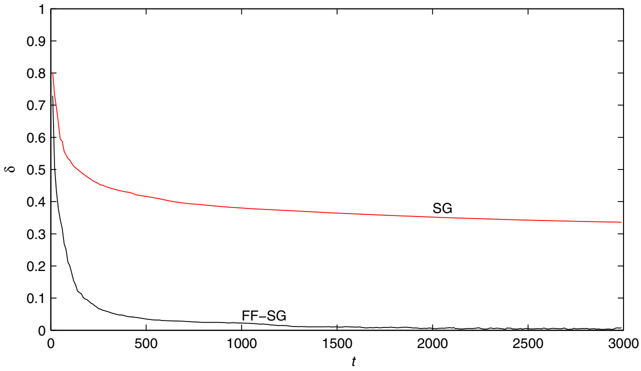

Fig. 2. The parameter estimation errors δ versus t

$$
\begin{aligned}
&= [\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5,]^{\mathrm{T}} \\
&= [0.2, 0.24, 0.9, 1.08, -0.1]^{\mathrm{T}}, \\
\boldsymbol{\varphi}(t) &= [h(u(t-1))u(t-1), h(u(t-2))u(t-2), \\
&\quad u(t-1), u(t-2), -y(t-1)]^{\mathrm{T}}.
\end{aligned}
$$

Applying the proposed SG and FF-SG algorithms to estimate the parameters of this system, the parameter estimates and their errors are shown in Tables 1-2 and the parameter estimation errors $\delta := \|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\|/\|\boldsymbol{\theta}\|$ versus t are shown in Figure 2.

Table 1. The SG estimates and errors

t	α_1	α_2	α_3	α_4	α_5	δ (%)
100	-0.0422	0.0043	0.4938	0.5448	-0.2357	52.9384
200	-0.0291	0.0224	0.5536	0.6047	-0.2483	47.3742
300	-0.0180	0.0347	0.5844	0.6351	-0.2537	44.4015
500	-0.0108	0.0442	0.6168	0.6669	-0.2658	41.6292
1000	0.0009	0.0534	0.6589	0.7046	-0.2646	37.9839
1500	0.0055	0.0585	0.6771	0.7217	-0.2663	36.4216
2000	0.0102	0.0630	0.6906	0.7342	-0.2663	35.2145
2500	0.0135	0.0666	0.7024	0.7442	-0.2663	34.2374
3000	0.0160	0.0690	0.7093	0.7510	-0.2653	33.5802
True values	0.2000	0.2400	0.9000	1.0800	-0.1000	

Let $\hat{\alpha}_i$ be the ith element of the vector $\hat{\boldsymbol{\theta}}$. From the definition of $\boldsymbol{\theta}$, we have: $\hat{a}_1 = \hat{\alpha}_5$, $\hat{b}_2 = \frac{\hat{\alpha}_2}{\hat{\alpha}_1}$. Furthermore, we can compute the estimates $\hat{m}_1 = \hat{\alpha}_3 + \frac{\hat{\alpha}_1}{2}$, $\hat{m}_2 = \hat{\alpha}_3 - \frac{\hat{\alpha}_1}{2}$.

Table 2. The FF-SG estimates and errors

t	α_1	α_2	α_3	α_4	α_5	δ (%)
100	0.1174	0.0619	0.8128	0.9035	-0.1812	20.0681
200	0.1927	0.1290	0.9055	1.0303	-0.1449	9.0010
300	0.2224	0.1689	0.8981	1.0440	-0.1097	5.7755
500	0.2285	0.1994	0.9009	1.0734	-0.1009	3.4647
1000	0.2206	0.2145	0.9003	1.0762	-0.1026	2.2923
1500	0.2105	0.2298	0.8962	1.0786	-0.1004	1.0502
2000	0.2033	0.2333	0.8983	1.0778	-0.1023	0.5736
2500	0.1952	0.2395	0.8952	1.0825	-0.0981	0.5179
3000	0.1980	0.2361	0.8958	1.0754	-0.1055	0.6487
True values	0.2000	0.2400	0.9000	1.0800	-0.1000	

5 Conclusions

An approach to identify Hammerstein systems with piece-wise linearity is presented in this paper. The model of the nonlinear system be turned into an identification model by using a switching function, then based on the identification model, we proposed an SG algorithm and an FF-SG algorithm to estimate all the parameters of the system. The simulation results verify the proposed algorithm.

References

1. Liu, Y., Bai, E.W.: Iterative identification of Hammerstein systems. Automatica 43(2), 346–354 (2007)
2. Ding, F., Liu, P.X., Liu, G.: Identification methods for Hammerstein nonlinear systems. Digital Signal Processing 21(2), 215–238 (2011)
3. Chen, J., Wang, X.P., Ding, R.F.: Gradient based estimation algorithm for Hammerstein systems with saturation and dead-zone nonlinearities. Applied Mathematical Modelling 36(1), 238–243 (2012)
4. Yu, L., Zhang, J.B., Liao, Y.W., Ding, J.: Parameter estimation error bounds for Hammerstein finite impulsive response models. Applied Mathematics and Computation 202(2), 472–480 (2008)
5. Wang, D.Q., Chu, Y.Y., Ding, F.: Auxiliary model-based RELS and MI-ELS algorithms for Hammerstein OEMA systems. Computers & Mathematics with Applications 59(9), 3092–3098 (2010)
6. Chen, J., Zhang, Y., Ding, R.F.: Auxiliary model based multi-innovation algorithms for multivariable nonlinear systems. Mathematical and Computer Modelling 52(9-10), 1428–1434 (2010)
7. Wang, D.Q., Chu, Y.Y., Ding, F.: Auxiliary model-based RELS and MI-ELS algorithms for Hammerstein OEMA systems. Computers & Mathematics with Applications 59(9), 3092–3098 (2010)
8. Ding, F., Shi, Y., Chen, T.: Auxiliary model-based least-squares identification methods for Hammerstein output-error systems. Systems & Control Letters 56(5), 373–380 (2007)

9. Bai, E.W.: Identification of linear systems with hard input nonlinearities of known structure. Automatica 38(5), 853–860 (2002)
10. Vörös, J.: Parameter identification of discontinuous Hammerstein systems. Automatica 33(6), 1141–1146 (1997)
11. Vörös, J.: Modeling and parameter identification of systems with multisegment piecewise-linear Characteristics. IEEE Transactions on Automatic Control 47(1), 184–188 (2002)
12. Vörös, J.: Parameter identification of Wiener systems with multisegment piecewise-linear nonlinearities. Systems & Control Letters 56(2), 99–105 (2007)
13. Jing, C., Lixing, L., Ruifeng, D.: Parameter Estimation for Dual-Rate Sampled Data Systems with Preload Nonlinearities. In: Zhang, T. (ed.) Mechanical Engineering and Technology. AISC, vol. 125, pp. 43–50. Springer, Heidelberg (2012)
14. Vörös, J.: Modeling and identification of systems with backlash. Automatica 46(2), 369–374 (2010)
15. Rochdi, Y., Giri, F., Gning, J.B., Chaoui, F.Z.: Identification of block-oriented systems in the presence of nonparametric input nonlinearities of switch and backlash types. Automatica 46(5), 864–877 (2010)