Global Parameter Estimation of Hammerstein Systems with Polynomial Nonlinearity

Cheng-ho CHEN
Department of Mechanical Engineering, National Chin-Yi University of Technology, Taiwan

Keywords: Hammerstein system, Parameter estimation, Nonlinear system, Polynomial nonlinearity.

Abstract. The parameter estimation problem of Hammerstein systems with polynomial nonlinearity is studied. An algorithm is proposed to overcome the problem of possible wrong convergence in Maximum Likelihood methods. The new algorithm provides a unique set of estimates for a given input and output data record corrupted by noise. The effectiveness and accuracy of the method are examined via numerical simulations.

Introduction
Proper modeling and identification of dynamical systems are essential in engineering. A Hammerstein-type system consists of a static nonlinearity followed by a linear dynamic subsystem. It has been widely studied in modeling nonlinear systems [1]-[7]. Various identification methods have been developed for the Hammerstein-type nonlinear systems. They can be categorized into parametric and non-parametric ones. Nonparametric methods [8]-[13] are able to capture the system characteristics usually with less computation. However, it is often desirable to have parametric models in engineering problems. Well developed methods for linear systems [14,15] can be applied to the Hammerstein system with polynomial type nonlinearity [16,17]. If the nonlinearity is smooth, other types of functions may be used to represent the nonlinearity [18]-[20]. Because of the nonlinear nature of the problem, appropriate initial values are required [21,22].

In this paper, a novel algorithm is proposed to obtain a unique set of estimates for the Hammerstein system parameters. The method applies linear regression and filtering operations so that possible erroneous convergence associated with nonlinear search schemes can be avoided. No particular types of the input signal are required. The system and noise dynamics along with nonlinear parameters are all estimated in a single model structure. Some simulation results are shown to demonstrate the effectiveness and accuracy of the method.

Problem Statement and Model Structure
The system to be considered in this paper, shown in Figure 1, is composed of a memoryless polynomial nonlinearity cascaded by a linear dynamic subsystem and has its output corrupted by a stationary stochastic noise process characterized by a rational spectral density.

Figure 1. A Hammerstein system with polynomial nonlinearity.
A formal mathematical representation of this system may be given as follows:

\[ S : \begin{cases} \dot{y}(t) = B^o(B) \cdot u(t) + C^o(B) \cdot w(t) \\ u(t) = \sum_{i=1}^{p} r_i \cdot x(t) \end{cases} \]  

where \( x(t), y(t), w(t) \) represents the input, output, and innovation signals, respectively. \( u(t) \) is the output of the static nonlinearity. \( A^o(B), B^o(B) \) and \( C^o(B) \) are functions in the back-shift operator \( B[Bx(t) = x(t-1)] \):

\[
A^o(B) \triangleq 1 + a_1^o B + a_2^o B^2 + \cdots + a_k^o B^k \\
B^o(B) \triangleq 1 + b_1^o B + b_2^o B^2 + \cdots + b_k^o B^k \\
C^o(B) \triangleq 1 + c_1^o B + c_2^o B^2 + \cdots + c_k^o B^k
\]

The superscript \( o \) is used to indicate characteristics associated with the true system \( S \) rather than any particular model. The following assumptions are made:

A1. \( A^o(B) \) and \( C^o(B) \) are strictly minimum phase.
A2. \( w(t) \) is zero-mean uncorrelated Gaussian, and uncross-correlated with \( x(t) \).

\[
E[w(t)]=0 \quad E[w(t) \cdot w(s)] = \delta_{t,s} \left( \sigma_w^2 \right) \quad E[w(t) \cdot x(s)] = 0 \quad \forall t,s
\]

where \( E[\cdot] \) denotes expectation, \( \delta_{t,s} \) the the Kronecker delta, and \( \left( \sigma_w^2 \right) \) the variance of \( w(t) \).

The problem may be stated as: “Given input and output observations \( x_1, \ldots, x(N) \) and \( y_1, \ldots, y(N) \), select a model \( M(\theta) \) that best matches the input-output and noise characteristics of the actual system \( S \).” The model structure \( M \) is selected as the set of all models \( M(\theta) \) of the form:

\[
M(\theta) : \begin{cases} \hat{y}(t/\theta) = \left[ 1 - \frac{A(B, \theta)}{C(B, \theta)} \right] y(t) + \frac{B(B, \theta)}{C(B, \theta)} \sum_{i=1}^{p} r_i(\theta) \cdot x(t) \\ e(t/\theta) = y(t) - \hat{y}(t/\theta) \end{cases}
\]

\( \theta \) represents the parameter vector to be estimated:

\[
\theta \triangleq [r_1, \ldots, r_p, \text{coef } A, \text{coef } B, \text{coef } C]^T
\]

To ensure identifiability, the following assumptions are added:

A3. \( A^o(B) \) and \( B^o(B) \), as well as \( A^i(B) \) and \( C^o(B) \), have no common factors.
A4. The leading coefficient of the \( B^o(B) \) function is equal to unity \( (b_0^o = 1) \).

The Parameter Estimation Method

An estimator based on Maximum Likelihood (ML) principle [23] and the aforementioned assumptions, can be formulated as:

\[
\hat{\theta} = \arg \min_{\theta} V_N(\theta) \triangleq \arg \min_{\theta} \frac{1}{N} \sum_{t=1}^{N} \frac{1}{2} e^2(t/\theta)
\]

The difficulty associated with this estimator stems from the non-quadratic nature of the minimization caused by the dependence of \( V_N(\theta) \) on \( C(B) \). To overcome this difficulty, define the following impulse response functions:
\[ H_y(B) A(B) = C(B) \]

\[ H_y(B) B(B) = C(B) \]

Truncated versions of the asymptotically stable sequences \( \{ h_{yi}^o \}, \{ h_{ui}^o \} \) along with the nonlinearity parameters \( \psi^o \), can be estimated within the model set:

\[ M_1(\theta_1): H_1(B / \theta_1) \cdot y(t) = H_u(B / \theta_1) \cdot [x(\theta_1)x'(\theta_1)] + e_1(t / \theta_1) \]

\[ H_y(B) A(B) = C(B) = 1 + h_{yi} B + \cdots + h_{yi_n} B^m \]

\[ H_u(B) B(B) = C(B) = B^{ak} \{ 1 + h_{ui} B + \cdots + h_{ui_n} B^m \} \]

\[ \theta_1 \Delta [h_{yi}, \ldots, h_{yi_n} ; r_1, \ldots, r_p ; h_{ui}, \ldots, h_{ui_n}] \]

\[ e_1(t / \theta_1) \] represents the one-step-ahead prediction error. The new model structure may be defined as:

\[ M_1(\theta_1): y(t) = \psi_1^T(t) \cdot \theta(\theta_1) + e_1(t / \theta_1) \]

Where

\[ \psi_1^T(t) \cdot \theta(\theta_1) \Delta \]

\[ \begin{bmatrix}
  - y(t - 1) \\
  \vdots \\
  - y(t - m) \\
  \vdots \\
  x(t - nk) \\
  \vdots \\
  x^p(t - nk) \\
  \vdots \\
  x(t - nk - 1) \\
  \vdots \\
  x(t - nk - m) \\
  \vdots \\
  x^p(t - nk - 1) \\
  \vdots \\
  x^p(t - nk - m)
\end{bmatrix} \begin{bmatrix}
  h_{yi} \\
  \vdots \\
  h_{yi_n} \\
  r_1 \\
  \vdots \\
  r_p \\
  r_1 h_{ui} \\
  \vdots \\
  r_p h_{ui_n}
\end{bmatrix} \]

\( \theta \) can be estimated by the following covariance-type LS estimator:
\[
\hat{\theta}_2 = \left( \frac{1}{N - n - n_k} \sum_{t=m+n_k+1}^N \psi_2(t) \cdot \psi_2^T(t) \right)^{-1} \cdot \left( \frac{1}{N - n - n_k} \sum_{t=m+n_k+1}^N \psi_2(t) \cdot y(t) \right)
\] (28)

where \( n2 \Delta \text{max}(na, nb, nc) + nc \). \( \epsilon^o \) can then be estimated as:

\[
\hat{\epsilon} = \hat{H}^{-1} \cdot \hat{\epsilon}
\] (22)

Afterwards, define:

\[
u_f(B) \Delta = \frac{\sum_{i=1}^p \hat{r}_i \cdot x_i(t)}{\hat{C}(B)} \quad y_f(B) \Delta = \frac{y(t)}{\hat{C}(B)}
\] (23)

The model in (6)-(7) can be then rewritten as:

\[
M_2(\theta_2) : A(B) \cdot \psi_2(t) = B(B) \cdot \psi_2(t) + \epsilon_2(t / \theta_2)
\] (24)

or equivalently,

\[
M_2(\theta_2) : \psi_2(t) = \psi_2^T(t) \cdot \theta_2 + \epsilon_2(t / \theta_2)
\] (25)

\[
\psi_2^T(t) \Delta = [-y_f(t-1), \ldots, -y_f(t-na), -u_f(t-1), \ldots, -u_f(t-nb)]
\] (26)

\[
\theta_2 \Delta = [a_{na}, b_{na}, \ldots, b_{nb}]
\] (27)

\( \theta_2 \) can then be estimated by the LS estimator:

\[
\hat{\theta}_2 = \left( \frac{1}{N - n - n_k} \sum_{t=m+n_k+1}^N \psi_2(t) \cdot \psi_2^T(t) \right)^{-1} \cdot \left( \frac{1}{N - n - n_k} \sum_{t=m+n_k+1}^N \psi_2(t) \cdot y_f(t) \right)
\] (28)

where \( n2 \Delta \text{max}(na, nb) \). \( A(B) \) and \( B(B) \) can be obtained from \( \theta_2 \). The proposed algorithm may be summarized in the following:
1. Obtain estimates of \( r_1, \ldots, r_p \) and truncated impulse responses functions \( \{h_{r_j}\}_{j=1}^m, \{h_{u_j}\}_{j=1}^m \) from the estimator (18).
2. Calculate \( \hat{C}(B) \) from Eq. (22).
3. Obtain \( \hat{A}(B) \) and \( \hat{B}(B) \) through Eq. (28).

**Simulation Results**

The parameters estimated from previous steps can be shown to be accurate with sufficiently large \( m \) and number of data points \( N \). In this section, the performance of the proposed method is evaluated via numerical simulations. The system output is generated by using mutually independent normal pseudo-random signals with zero mean as excitation and innovations sequences. The noise-to-signal \((N/S)\) ratio is defined as:

\[
N/S = \frac{\text{Var}[w(t)]}{\text{Var}[x(t)]} \times 100\%
\]

Equation (29)

\( p^o, na^o, nb^o, nk^o \) and \( ne^o \) are assumed known a-priori. Table 1 presents a Monte Carlo analysis of the method based on 20 data records with \( N/S=30\% \). Each data record consists of 2000 points for estimation. The order of truncated impulse response functions, \( m \), is set to be 20. The results show that the estimates are very accurate.

| System Parameters | Estimated Parameters |
|-------------------|----------------------|
| \( r_1 \)        | 0.300 0.298 0.0052  |
| \( r_2 \)        | 0.500 0.500 0.0024  |
| \( r_3 \)        | 0.800 0.800 0.0032  |
| \( a_1 \)        | -1.650 -1.650 0.0016 |
| \( a_2 \)        | 0.750 0.750 0.0014  |
| \( b_1 \)        | 0.700 0.700 0.0074  |
| \( b_2 \)        | -0.400 -0.400 0.0074 |
| \( c_1 \)        | -1.200 -1.198 0.0017 |
| \( c_2 \)        | 0.600 0.598 0.0028  |

Number of data sets: 20. \( N/S = 30\% \)

**References**

[1] P.H. Menold, F. Allgower, Pearson, R.K. Nonlinear structure identification of chemical processes. Computers chem. Eng. 21 (1997), Suppl., s137-s142.

[2] T. Kara, İ Eker, Nonlinear modeling and identification of a DC motor for bidirectional operation with real time experiments. Energy Conversion and Management. 45 (2004), 1087-1106.

[3] M. Milanese, C. Novara, Structure set membership identification of nonlinear Systems with application to vehicles with controlled suspension. Control Engineering Practice 15 (2007), 1-16.

[4] D. Rollins, N. Bhandar, S. Hulting, System Identification of the Human Thermoregulatory System Using Continuous-Time Block-oriented predictive modeling. Chemical Engineering Science 61 (2006) 1516-1527.
[5] K.-J. Xu, H. Ren, X.-F. Wang, Q. Teng, Non-linear dynamic modeling of hot-film/Wire MAF sensors with two-stage identification based on Hammerstein model. Sensors and Actuators A 135 (2007) 131-140.

[6] F. Le, I. Markovsky, C.T. Freeman, E.Rogers, Recursive identification of Hammerstein systems with application to electrically stimulated muscle, Control Engineering Practice 20 (2012) 386-396.

[7] C.M. Holcomb, R.A. de Callafon, R.R. Bitmead, Closed-Loop Identification of Hammerstein Systems with Application to Gas Turbines, Proceedings of the 19th World Congress The International Federation of Automatic Control (2014) 493-498.

[8] A. Kryzyżak, On nonparametric estimation of nonlinear dynamic systems by the Fourier series estimate. Signal Processing. 52 (1996) 299-321.

[9] W. Greblicki, Nonlinearity estimation in Hammerstein systems based on ordered observations. IEEE Transactions on Signal Processing 44 (1996) No.5, 1224-1233.

[10] Z. Hasiewicz, Non-parametric estimation of non-linearity in a cascade time-series system by multiscale approximation, Signal Processing 81 (2001) 791-807.

[11] A. Kryzyżak, M.A. Partyka, Global identification of nonlinear Hammerstein systems by cascade time-series system by recursive kernel approach, Nonlinear Analysis 63 (2005) e1263-e1272.

[12] P. Hasiewicz, Recursive estimation of Hammerstein system nonlinearity by Haar wavelets, 8th IFAC Symposium on Nonlinear Control Systems 2010, 439-444.

[13] R.S. Risuleo, G. Bottegal, H. Hjalmarsson, A kernel-based approach to Hammerstein system identification, IFAC-PapersOnLine 48-28 (2015) 1011–1016.

[14] L. Ljung, System identification: theory for the user, Prentice Hall, Englewood Cliffs, N.J., USA, 1987.

[15] T. Söderström, P. Stoica, System identification, Prentice Hall, Cambridge, U.K. 1989.

[16] T. Kara, İ. Eker, Experimental nonlinear identification of a two mass system, Proceedings of IEEE Conference on Control Applications 1 (2003) 66-71.

[17] F. Ding, T. Chen, Identification of Hammerstein nonlinear ARMAX systems, Automatica 41 (2005) 1479-1489.

[18] J.C. Gómez, E. Baeyens, Identification of block-oriented nonlinear system using orthonormal bases, Journal of Process Control 14 (2004) 685-697.

[19] K.H. Chan, J. Bao, W.J. Whiten, Identification of MIMO Hammerstein systems using cardinal spline functions, Journal of Process Control 16 (2006) 659-670.

[20] J. Li, Parameter estimation for Hammerstein CARARMA systems based on the Newton iteration, Applied Mathematics Letters 26 (2013) 91–96.

[21] P. Crama, P.J. Schoukens, R. Pintelon, Generation of enhanced initial estimates for Hammerstein Systems, Automatica 40 (2004) 1269-1273.

[22] J. Schoukens, W.D. Widanage, K. R. Godfrey, R. Pintelon, Initial estimates for the dynamics of Hammerstein system, Automatica 43 (2007), 1296-1301.

[23] K.J. Åström, Maximum likelihood and prediction error methods, Automatica 16 (1980) 551-574.