Identification of Hammerstein Models to minimize the H_∞ Norm of the Mismatch Error

Samir H. Al-Amer and Foaud M. AL-Sunni

samir@ccse.kfupm.edu.sa alsunni@ccse.kfupm.edu.sa

Systems Engineering Department

King Fahd University of Petrol. & Minerals

Dhahran 31261 Saudi Arabia

May 29, 2002

Abstract

In this paper we consider identification of a special class of non-linear systems using Hammerstein models. Hammerstein Models are special in the sense that they can be transformed into linear models and linear systems techniques may be applicable. An iterative identification algorithm is proposed to obtain a model that minimizes the infinity norm of the mismatch error. Illustrative examples are given.

1 Introduction

Linear models are often used to model real systems. They are easy to work with and the theory is well developed. In many cases, however, linear models are not adequate to represent the physical system and nonlinear models are needed. General nonlinear models may require considerable amount of computation to identify and to analyze. Hammerstein models have been successively used in modeling some physical systems [1]. The Hammerstein

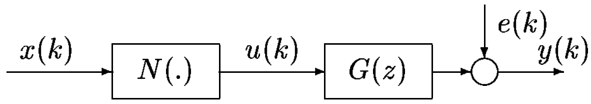

Figure 1: Hammerstein Model

model consists of a memoryless nonlinear subsystem $N(.)$ followed by a linear shift invariant subsystem $G(z)$ as shown in Figure 1. The intermediate variable $u(k)$ is not measurable.

Several algorithms to identify Hammerstein models have been proposed They can be classified into two groups [2, 3]. In the first class iterative identification of the linear and nonlinear parts is done [4, 5]. A major problem in these algorithms is assuring the convergence of the iterations[6]. The second group includes noniterative algorithms that simultaneously identify $N(.)$ and $G(.)$ [7]. In general, these are expected to lead to more accurate models but they have a larger number of parameters to be estimated.

The least squares and mean square error techniques are often used to obtain the models. An algorithm to minimize the ℓ_∞ gain is presented in [8].

In this paper we propose an identification algorithm to obtain a Hammerstein model that minimizes the infinity norm of the mismatch error. The rest of the paper is organized as follows: Section 2 presents the problem statement and the notation used. The identification algorithm is given in Section 3 and numerical examples are given in Section 4.

2 Problem Statement

Given a set of input-output data $\{x(k), y(k)\}$ and two integers m and n, it is required to fit the data to a Hammerstein model. The system to be identified is a nonlinear single-input-single-output system. The nonlinear block is assumed to be of the form

$$u(k) = \sum_{i=1}^{m} v_i \ x^i(k) \tag{1}$$

and the linear block is assumed to be a stable shift invariant system described by

$$y(k) = \frac{\sum_{i=0}^{n} b_i q^{-i}}{1 + \sum_{i=1}^{n} a_i q^{-i}} u(k) + e(k). \tag{2}$$

The sensor noise $e(k)$ is assumed to bounded. Let the model parameters be denoted by

$$\theta_v = [v_1 \ v_2 \ \cdots \ v_m]^T$$

$$\theta_a = [a_1 \ a_2 \ \cdots \ a_n]^T$$

$$\theta_b = [b_0 \ b_1 \ \cdots \ b_n]^T$$

Note that the parameters of the model can not be uniquely determined and the following constraint is introduced.

$$v_1 = 1.$$

The other coefficients are normalized accordingly. By defining the following artificial inputs [9]

$$z_i(k) = x^i(k),$$

the Hammerstein model can be viewed as a linear m-input 1-output system (see Figure 2 for illustration).

Let $Z_i(z)$, $U(z)$,$Y(z)$ and $E(z)$ denote the Z-transform of $z_i(k)$,$u(k)$,$y(k)$ and $e(k)$ respectively then

$$Y(z) = \frac{\sum_{i=0}^{n} b_i z^{-i}}{1 + \sum_{i=1}^{n} a_i z^{-i}} \times \sum_{i=1}^{m} v_i \, Z_i(z) + E(z) \tag{3}$$

In this paper we will present an identification algorithm to obtain v_i, a_i and b_i so that the H_∞ norm of the mismatch error is minimized. More precisely the following problem is to be solved.

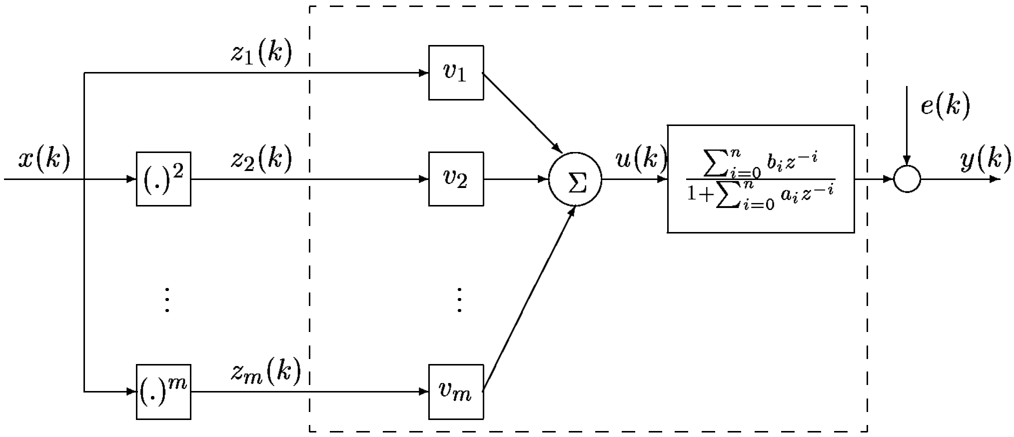

Figure 2: An Alternative Representation of Hammerstein Model

$$\min_{a_i, b_i, v_i} \left\| Y(z) - \frac{\sum_{i=0}^{n} b_i z^{-i}}{1 + \sum_{i=1}^{n} a_i z^{-i}} \sum_{i=1}^{m} v_i \, Z_i(z) \right\|_\infty \tag{4}$$

3 The Proposed Identification Algorithm

In this section we present a procedure to identify a Hammerstein model from a set of input output data to solve the optimization problem (4). The proposed algorithm is an extension of the algorithm developed in [10, 11]. The solution to the H_∞ approximation problem is obtained by solving a sequence of weighted least squares problems. Each iteration of the algorithm involves two approximation problems. In the first, θ_b is assumed to be available and a least squares problem is solved for θ_v and θ_a. In the second θ_v is fixed to its latest value and a least squares problem is solved for θ_a and θ_b. To insure the uniqueness of the parameter set, a scaling is done to force v_1 to be one and the other parameters are adjusted accordingly.

3.1 Implementation

The algorithm uses samples of the of the Z-transform of the input and output on the unit circle. This is obtained by computing the N-point discrete Fourier transform of $y(k)$ and $z_i(k)$. The computed Fourier transforms denoted by $Y(e^{j\omega_k})$ and $Z_i(e^{j\omega_k})$ are basically samples of $Y(z)$ and $Z_i(z)$ on the unit circle at the frequencies $\omega_k = \frac{2\pi k}{N}$.

Define

$$\phi_v(k) = \left[Z_1(e^{j\omega_k}),\ Z_2(e^{j\omega_k}), \cdots,\ Z_m(e^{j\omega_k})\right]$$

$$\phi_a(k) = \left[Y(e^{j\omega_k})e^{-j\omega_k},\ Y(e^{j\omega_k})e^{-j2\omega_k},\ \cdots,\ Y(e^{j\omega_k})e^{-jn\omega_k}\right]$$

$$\phi_b(k) = \left[1,\ e^{-j\omega_k},\ \cdots,\ e^{-jn\omega_k}\right]$$

$$\overline{\phi}_b(k) = \phi_v(k)\theta_v\phi_b(k)$$

and

$$\Phi_v = \begin{bmatrix} \phi_v(1) \\ \phi_v(2) \\ \cdots \\ \phi_v(N) \end{bmatrix} \quad \Phi_a = \begin{bmatrix} \phi_a(1) \\ \phi_a(2) \\ \cdots \\ \phi_a(N) \end{bmatrix} \quad \Phi_b = \begin{bmatrix} \overline{\phi}_b(1) \\ \overline{\phi}_b(2) \\ \cdots \\ \overline{\phi}_b(N) \end{bmatrix}$$

Each iteration of the propose algorithm involves solving two weighted least squares approximation problems. In the first problem, θ_b is fixed to its latest computed value and the following least square problem is solved for θ_v and θ_a.

$$\min_{\theta_v,\, \theta_a} \left\| \left(Y(e^{j\omega_k}) + \phi_a(k)\theta_a - \phi_b(k)\theta_b\ \phi_v(k)\theta_v\right) W(e^{j\omega_k}) \right\|_2 \tag{5}$$

In the second problem, θ_v is fixed at at its latest value and the following problem is solved for θ_a and θ_b.

$$\min_{\theta_a,\, \theta_b} \left\| \left(Y(e^{j\omega_k}) + \phi_a(k)\theta_a - \phi_b(k)\theta_b\ \phi_v(k)\theta_v\right) W(e^{j\omega_k}) \right\|_2$$

An initial guess for the linear part in the first problem is to assume that the linear block is an all-pole system (i.e., $b_0 = 1, b_i = 0$ for $i \neq 0$). The frequency weighting at each iteration is computed as the product of the previous weight and the error computed using the latest available model. An initial value of the weight is given by $W(e^{j\omega_k}) = 1 \ \forall k$. A summary of the proposed algorithm is given below.

The identification Algorithm

Given m, n, N, and $\{(x(k), y(k)), k = 1, 2, ...\}$.

Step 0:Compute the N-point FFT of $y(k)$ and $z_i(k)$.

Step 1: Let $b_0 = 1, b_i = 0$ for $i \neq 0$, $l = 1$,$W_k^l = 1$ for $k \in [1, N]$.

Step 2: Compute θ_a and θ_v using

$$\begin{bmatrix} \theta_a \\ \theta_v \end{bmatrix} = (\Phi^T W \Phi)^{-1} \Phi^T W Y$$

where

$$\Phi = [-\Phi_a \quad -\Phi_v]$$

and

$$W = diag(W^l(1) \ W^l(2), \cdots, W^l(N))$$

Step 3: Scale θ_b and θ_v using

$$\theta_b = v_1 \ \theta_b, \qquad \theta_v = (\frac{\theta_v}{v_1})$$

Step 4: Update the weight using

$$W^l(k) = W^l(k) \times \left| Y(k) - \frac{\phi_b(k)\theta_b}{1 + \phi_a(k)\theta_a} \phi_v(k)\theta_v \right|$$

Step 5: Now fix θ_v and Compute θ_b and θ_a using

$$\begin{bmatrix} \theta_a \\ \theta_b \end{bmatrix} = (\Phi_2^T \ W \ \Phi_2)^{-1} \Phi_2^T W Y$$

where

$$\Phi_2 = [-\Phi_a \quad \Phi_b]$$

Step 6: Update the weight as in step 4.

Step 7: Set l=l+1 and goto step 2.

Remark 1 *The algorithm is terminated after a fixed number of iteration and the identified model is selected as the one that gives the least error.*

4 Examples

Two examples are presented here to illustrate the algorithm.

4.1 Example 1

The nonlinear system being identified is shown in Figure 3. The apriori information $n = 1, m = 2$ are assumed

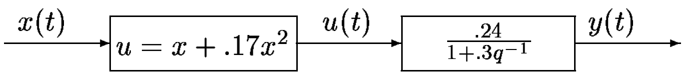

Figure 3: True Model, Example 1

The input x(k) is generated as a uniformly distributed sequence of magnitude 1. The measurement noise is uniformly distributed with different levels l=0,.02 and 0.05. The algorithm uses 1024-point FFT for computing the identified model.

The identified model(after 10 iterations) is given in Table 1.

Noise Level	Nonlinear block	Linear Block	$\|Error\|_\infty$
0	$u(k) = x(k) + 0.1698x^2(k)$	$\hat{G}(z) = \frac{0.2400}{z+0.2999}$	0.0126
.01	$u(k) = x(k) + 0.1742x^2(k)$	$\hat{G}(z) = \frac{0.0002z+0.2392}{z+0.3002}$	0.5004
.05	$u(k) = x(k) + 0.1391x^2(k)$	$\hat{G}(z) = \frac{-0.0012z+0.2416}{z+0.2759}$	2.1926

Table 1: Identified Models With different measurement noise level

It is observed that the infinity norm of the error is close to the maximum magnitude of the frequency response of the error signal.

4.2 Example 2

Consider the following nonlinear system. The nonlinear block is described by

$$x(t) = x(t) + 0.27x^2(t) + 0.15x^3(t)$$

and the linear part is described by

$$G(z) = \frac{0.6 + 0.25z^{-1} + 0.22z^{-2} + 0.1z^{-3}}{1 + 0.5z^{-1} + 0.3z^{-2} + 0.1z^{-3}}$$

Sequence of 300 input-output data points are generated with measurement noise levels $0, 0.01$ and 0.05. Ten iterations of the proposed algorithm was used with $N = 1024$ and the result is shown in Table 2.

Noise	Nonlinear block	Linear Block	$\|\|Error\|\|_\infty$
0	$u = x + 0.2723x^2 + 0.1488x^3$	$\hat{G}(z) = \frac{0.5996+0.2259z^{-1}+0.2241z^{-2}+0.0893z^{-3}}{1+0.4599z^{-1}+0.3039z^{-2}+0.0872z^{-3}}$	0.0303
0.01	$u = x + 0.2889x^2 + 0.1383x^3$	$\hat{G}(z) = \frac{0.5953+0.0327z^{-1}+0.2381z^{-2}+0.0103z^{-3}}{1+0.1368z^{-1}+0.3031z^{-2}+0.0088z^{-3}}$	0.2306
0.05	$u = x + 0.5836x^2 + 0.0370x^3$	$\hat{G}(z) = \frac{0.5199+0.0165z^{-1}+0.1249z^{-2}+0.0065z^{-3}}{1+0.1036z^{-1}+0.1380z^{-2}-0.0289z^{-3}}$	1.1529

Table 2: Identified Models With different measurement noise level

5 Conclusions

In this paper, a new iterative procedure to identify Hammerstein models is proposed. The algorithm minimizes the infinity norm of the deviation between the true model and identified model. Illustrative examples were given to demonstrate the algorithm. It is observed form the above examples that the infinity norm of the deviation error is close to the maximum magnitude of the frequency response of the measurement noise.

Acknowledgment

The authors acknowledge King Fahd University of Petroleum and Minerals for its support.

References

[1] E. Eskinat, S.H. Johnson and L. Luyben (1991), " Use of Hammerstein models in identification of nonlinear system," Am. Inst. Chem. Eng. J., Vol. 37, pp. 255-268.

[2] S. A. Billings (1980)," Identification of Nonlinear systems-A survey", IEE Proc., Part.D, Vol. 127, No 6, pp. 272-285, .

[3] R. Haber and H. Unbehauen (1990), " Structure identification of Nonlinear Dynamic Systems-A survey on input-output Approach", Automatica, vol 26, no 4, pp. 464-468.

[4] K.S. Narendra and P.G. Gallman (1966), "An iterative method for the Identification of Nonlinear Systems Using a Hammerstein Model", IEEE Trans. Auto. Control AC-11, No3, pp. 546-550.

[5] J. Voros,(1999), " Iterative Algorithm for parameter Identification of Hammerstein Systems with Two-Segment Nonlinearity", AC-44, No 11, pp. 2145-2149.

[6] P. Stoica (1981), " On The Convergence of an Iterative Algorithm Used for Hammerstein System Identification", IEEE Trans. Auto. Control Vol AC-26, pp967-969.

[7] F.H.I Chang and R. Luss (1973), "A noniterative method for identification using Hammerstein model",IEEE Trans AC-16, pp 464-468.

[8] H. Fukushima and T. Sugie(2001), "Worst-Case Identification of Hammerstein Models Based on ℓ_∞ Gain", Proc. 2001 ACC, Arlignton, VA, pp. 5022-5027.

[9] Soderstrom T. and P. Stoica (1989), *Systems Identification*, Prentice Hall International, U.K.

[10] D. Kavranğlue and S.H. Al-Amer (1996), " A New Frequency Domain Technique for H_∞ Norm Approximation", Proc. CDC 96, Kobe, Japan , pp. 4574-4579.

[11] D. Kavranğlue and S.H. Al-Amer (2001), " New Efficient Frequency Domain Algorithm for H_∞ Approximation with Applications to Controller Reduction", IEE Proc. Vol 148, Pt. D. pp. 383-390.