Iterative Identification of Continuous Hammerstein and Wiener Models

Ho-Tsen Chen (陳厚岑), Shyh-Hong Hwang (黃世宏), Chuei-Tin Chang (張玨庭)
Department of Chemical Engineering
National Cheng Kung University
Tainan 70101, Taiwan, R.O.C.

Abstract

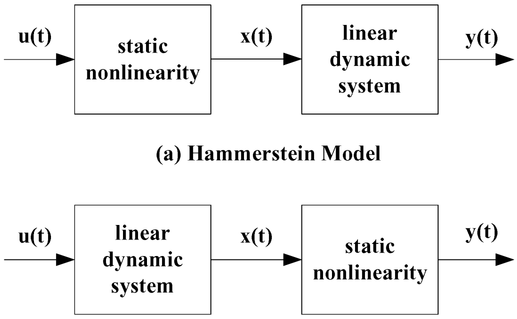
This paper presents iterative estimation algorithms to deal with parameter estimation of continuous Hammerstein and Wiener models. These two simple classes of blocked-oriented nonlinear systems consist of a nonlinear static block and a linear dynamic block in cascade. The internal variable between the two blocks is inaccessible to measurement so that the model parameters cannot be estimated by the conventional least-squares approach merely based on input and output measurements. To overcome this difficulty, the proposed algorithms are started by an initial guess of the internal variable and the resulting parameter estimates converge rather fast to their accurate values in an iterative manner. The use of a time-weighted integral transform can eliminate time derivatives of the variables in the model equation and renders the algorithms robust with respect to noise and model structure mismatch. Moreover, it ensures the convergence and accuracy of the algorithms to a great extent.

1. Introduction

Most of chemical processes are continuous and nonlinear. Identifying such processes as linear ones is often limited to a narrow range of operation. A remedy is to assume block-oriented nonlinear models such as Hammerstein and Wiener types [1]. Two such examples are a valve characteristic with a linear dynamic process and a chemical reaction followed by measurement of pH [2].

A Hammerstein model consists of a nonlinear static element followed by a linear dynamic system, whereas a Wiener model is constructed by a linear dynamic system followed by a nonlinear static element. Huang et al. [3] and Lee and Huang [4] proposed to identify a continuous Hammerstein model and a continuous Wiener model, respectively, using relay feedback experiments. Their methods employ an optimization procedure to find the inverted nonlinear function that restores symmetric cycling of the output of the relay system. With the nonlinear function given, the internal variable can be calculated and the linear dynamic part can then be identified using any available linear technique. Some limitations of these methods are that the nonlinear static element must be monotonic and the computation burden is quite heavy.

Voros [5] developed an iterative scheme to identify a discrete Hammerstein model. The same idea was later extended to identification of a Wiener model [6]. In his approach, the nonlinear system is modeled as a linear difference equation in conjunction with a nonlinear polynomial function. The entire equation is arranged so that the conventional prediction-error method is applicable to parameter estimation provided that the internal variable is given. The iteration procedure is then started by an initial guess of the internal variable. At each iteration step, the internal variable is updated using a recursive formula involving current estimates of the parameters and the internal variable. The major disadvantage of Voros's approach is that the convergence and accuracy of parameter estimates are not warranted especially when noise is present or the model structure (orders and delay) is not correct.



(b) Wiener Model

Figure 1. Two classes of block-oriented nonlinear systems.

In this work, we apply Voros's idea on discrete iterative estimation to identify a continuous Hammerstein or Wiener model. A time-weighted integral transform is incorporated into the iterative estimation algorithms to enhance the convergence and accuracy. Furthermore, the resulting method is robust with respect to noise and model structure mismatch.

2. Nonlinear Systems Description

Here, we consider continuous identification of two quadratic block-oriented nonlinear systems, i.e., the Hammerstein model and the Wiener model. The two models are characterized by the cascade connection of a linear dynamic block and a nonlinear static block as depicted in Fig. 1. In the Hammerstein model, the nonlinear static block receives the input signal $u(t)$ and sends the converted signal to the linear dynamic block. In the Wiener model, the nonlinear static block receives the signal from the linear dynamic block and generates the output signal $y(t)$. The signal between the linear dynamic block and the nonlinear static block is called the internal variable $x(t)$. It is assumed that the input and output variables are measurable whereas the internal variable is inaccessible to measurement.

Because the internal variable $x(t)$ is unknown, the conventional least-squares method for linear-in-parameter estimation is not applicable. Obviously, the resulting linear regression equation must consist of combinations of parameters and the unknown internal variable. This difficulty can be overcome by estimating the model parameters with an assumed estimate of the internal variable and updating it in an iterative fashion as will be elaborated later.

3. Identification of Hammerstein Models

The dynamic part of the Hammerstein model relating the internal variable $x(t)$ to the output variable $y(t)$ can be described by the following linear system equation:

$$\begin{aligned} &a_n y^{(n)}(t) + a_{n-1} y^{(n-1)}(t) + \cdots + a_1 y^{(1)}(t) \\ &\quad + y(t) = b_m x^{(m)}(t-d) + b_{m-1} x^{(m-1)}(t-d) \\ &\quad + \ldots + b_1 x^{(1)}(t-d) + b_0 x(t-d) \end{aligned} \quad (1)$$

where a_i and b_i are model parameters, n and m denote the system orders, and d is time delay. Moreover, we assume that the static part relating the input variable $u(t)$ to the internal variable $x(t)$ can be approximated by a polynomial:

$$x(t) = \sum_{j=1}^{p} c_j \left[u(t)\right]^j \quad (2)$$

where p denotes the order of the polynomial and c_i are also model parameters. Without loss of generality and from the viewpoint of the input-output relationship, the parameter b_0 can be arbitrarily assigned to be 1 and allow c_i to account for the gain of the model.

Substituting Eq. (2) into the $x(t-d)$ term of Eq. (1) and rearranging the resulting equation gives rise to

$$\begin{aligned} y(t) &= -\sum_{j=1}^{n} a_j y^{(j)}(t) \\ &\quad + \sum_{j=1}^{m} b_j x^{(j)}(t-d) + \sum_{j=1}^{p} c_j \left[u(t-d)\right]^j \end{aligned} \quad (3)$$

Equation (3) constitutes a linear regression equation provided that the variables $y(t)$, $u(t)$, and $x(t)$ are given. However, there are two difficulties involved in least-squares parameter estimation based on this equation. First, time derivatives of these variables are not available. Second, the internal variable $x(t)$ is not measurable. The first difficulty can be eliminated by applying an integral transform on the equation, while the second difficulty can be overcome by performing the least-squares estimation procedure in an iterative manner.

4. Identification of Wiener Models

The Wiener model is described by

$$\begin{aligned} &a_n x^{(n)}(t) + a_{n-1}x^{(n-1)}(t) + \cdots + a_1 x^{(1)}(t) + x(t) = \\ &b_m u^{(m)}(t-d) + b_{m-1}u^{(m-1)}(t-d) + \cdots \\ &+ b_1 u^{(1)}(t-d) + b_0 u(t-d) \end{aligned} \tag{4}$$

and

$$y(t) = \sum_{j=1}^{p} c_j \left[x(t)\right]^j \tag{5}$$

Equation (4) relates the input variable $u(t)$ to the internal variable $x(t)$ whereas Eq. (5) relates the internal variable $x(t)$ to the output variable $y(t)$. Without loss of generality, we let $c_1 = 1$ and obtain the following expression:

$$x(t) = y(t) - \sum_{j=2}^{p} c_j \left[x(t)\right]^j \tag{6}$$

Substituting Eq. (6) into the $x(t)$ term in Eq. (4) yields

$$\begin{aligned} y(t) = &-\sum_{j=1}^{n} a_j x^{(j)}(t) \\ &+ \sum_{j=0}^{m} b_j u^{(j)}(t-d) + \sum_{j=2}^{p} c_j \left[x(t)\right]^j \end{aligned} \tag{7}$$

Equation (7) constitutes a linear regression equation for the Wiener model and results in the same difficulties encountered in least-squares parameter estimation of the Hammerstein model.

5. Time-Weighted Integral Transform

To deal with the time derivatives in Eq. (3) and (7), we employ the time-weighted integral transform proposed by Hwang and Lin [7]. The ith-order integral transform is to convert a continuous signal $f(t)$ over the time interval $[t_a, t_b]$ into a real number:

$$T_i\{f(t)\} = F_i(t_a, t_b) = \int_{t_a}^{t_b} w^{(i)}(t) f(t)\, dt \tag{8}$$

where the superscript (i) denotes the ith-order derivative of the weighting function $w(t)$ with respect to time. The zeroth-order transform of $f^{(i)}(t)$ can be derived as the following form:

$$\begin{aligned} T_0\{f^{(i)}(t)\} &= \int_{t_a}^{t_b} w(t) f^{(i)}(t)\, dt \\ &= (-1)^i T_i\{f(t)\} + \sum_{j=0}^{i-1} (-1)^j \times \\ &\left[w^{(j)}(t_b) f^{(i-1-j)}(t_b) - w^{(j)}(t_a) f^{(i-1-j)}(t_a)\right] \end{aligned} \tag{9}$$

Suppose that the following weighting function is proposed:

$$w(t) = (t - t_a)^n (t - t_b)^n \tag{10}$$

It is apparent that for $i \le n$

$$w^{(i-1)}(t_a) = w^{(i-2)}(t_a) = \cdots = w(t_a) = 0$$
$$w^{(i-1)}(t_b) = w^{(i-2)}(t_b) = \cdots = w(t_b) = 0$$

Then all initial and final states of the signal in Eq. (9) can be eliminated. As a result, it reduces to

$$T_0\{f^{(i)}(t)\} = (-1)^i T_i\{f(t)\} = (-1)^i F_i(t_a, t_b) \tag{11}$$

Taking the zeroth-order transform on both sides of the Eq. (3) and applying Eq. (10) gives rise to the new regression equation for the Hammerstein model:

$$Y_0(t_a, t_b) = \phi_h(t_a, t_b)^T \theta_h \tag{12}$$

where

$$\phi_h = \begin{bmatrix} (-1)^{n-1} Y_n \\ (-1)^{n-2} Y_{n-1} \\ \vdots \\ Y_1 \\ (-1)^m XD_m \\ (-1)^{m-1} XD_{m-1} \\ \vdots \\ -XD_1 \\ UD_0^1 \\ UD_0^2 \\ \vdots \\ UD_0^p \end{bmatrix}, \quad \theta_h = \begin{bmatrix} a_n \\ a_{n-1} \\ \vdots \\ a_1 \\ b_m \\ b_{m-1} \\ \vdots \\ b_1 \\ c_1 \\ c_2 \\ \vdots \\ c_p \end{bmatrix}$$

$$Y_i(t_a,t_b)=\int_{t_a}^{t_b} w^{(i)}(t)y(t)\,dt$$

$$XD_i(t_a,t_b)=\int_{t_a}^{t_b} w^{(i)}(t)x(t-d)\,dt$$

$$UD_0^i(t_a,t_b)=\int_{t_a}^{t_b} w(t)[u(t-d)]^i\,dt$$

Similarly, we obtain the regression equation for the Wiener type from Eq. (7) as

$$Y_0(t_a,t_b)=\phi_w(t_a,t_b)^T\,\theta_w \tag{13}$$

where

$$\phi_w=\begin{bmatrix}(-1)^{n-1}X_n\\(-1)^{n-2}X_{n-1}\\\vdots\\X_1\\(-1)^m UD_m\\(-1)^{m-1}UD_{m-1}\\\vdots\\UD_0\\X_0^2\\X_0^3\\\vdots\\X_0^p\end{bmatrix},\quad \theta_w=\begin{bmatrix}a_n\\a_{n-1}\\\vdots\\a_1\\b_m\\b_{m-1}\\\vdots\\b_0\\c_2\\c_3\\\vdots\\c_p\end{bmatrix}$$

$$X_i(t_a,t_b)=\int_{t_a}^{t_b} w^{(i)}(t)x(t)\,dt$$

$$UD_i(t_a,t_b)=\int_{t_a}^{t_b} w^{(i)}(t)u(t-d)\,dt$$

$$X_0^i(t_a,t_b)=\int_{t_a}^{t_b} w(t)[x(t)]^i\,dt$$

A simple approach for least-squares parameter estimation is to use Eq. (12) or (13) to generate a large number ($N >> m+n+p$) of linear regression relations by choosing different time intervals (or horizons) for integration. We recommend choosing

$$t_a(k)=d+0.1(k-1)\lambda$$
$$t_b(k)=t_a(k)+\lambda;\qquad k=1,2,\cdots,N \tag{14}$$

Each time horizon starts from a different $t_a(k)$ and has the same length λ.

6. Iterative Estimation Procedure

The remaining problem in Eqs. (12) and (13) is that the internal variable is unknown so that the conventional least-squares parameter estimation method is not appropriate. This problem can be circumvented by the iterative estimation procedure described as follows. First construct Eq. (12) for the Hammerstein model or Eq. (13) for the Wiener model with a guess of the internal variable. A convenient guess is the measured input variable for the Hammerstein model and the measured output variable for the Wiener model. This in conjunction with Eq. (14) gives rise to the first estimates of the model parameters. The internal variable can now be updated using Eq. (2) for the Hammerstein model and Eq. (6) for the Wiener model. The preceding step is then repeated to give a new set of parameter estimates. This iterative procedure is continued until the parameter estimates converge.

Two issues arise, i.e. convergence and accuracy of the iterative algorithms. For the iterative algorithms being useful, the parameter estimates need to converge to the accurate values. The advantage of our algorithms is that the convergence and accuracy can be greatly improved by using a large value of the estimation horizon λ. A plausible reason is that if λ is sufficiently large, at each iteration the algorithms utilize a sufficient amount of information over the time horizon λ about the system, thereby ensuring convergence and accuracy.

7. Simulation Examples

Three examples are employed to evaluate the proposed iterative algorithms. The λ values of the three examples are chosen to be 15, 20, and 5. The model orders and delays are assumed given (not necessarily correct). The order of the polynomial for each example is arbitrarily set to 3. The input signal for identification is a white random signal varied at time instants kT ($T=1$ and $k=0,1,2,\cdots$). The magnitude of the input signal should be large enough to excite the nonlinear static property of the system. To test the robustness of the algorithms, measurement noise is added to the output variable for the first two examples with the noise-to-signal ratio (NSR) defined as the standard deviation of the noise divided by the standard deviation of the output signal. The third example is under noise-free conditions.

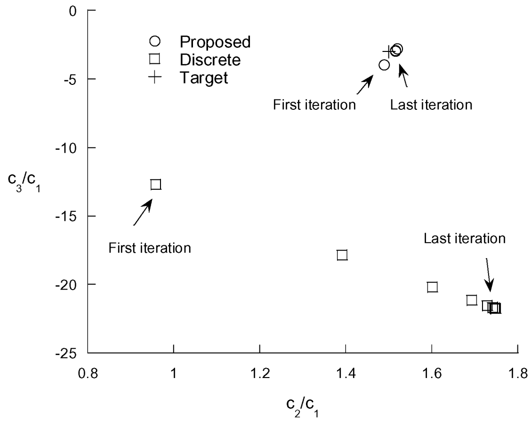


Figure 2. Comparison of the coefficients of the polynomial estimated in the iterative procedures of example 1.

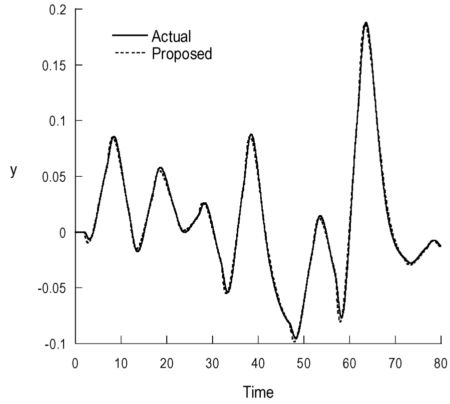


Figure 3. Comparison of the actual response of example 1 with the model predictions obtained by the proposed algorithms with reduced order.

Example 1

$$2y^{(3)}(t)+5y^{(2)}(t)+4y^{(1)}(t)+y(t)= -x^{(1)}(t-2)+x(t-2)$$

$$x(t)=u(t)+1.5[u(t)]^2-3[u(t)]^3$$

This example is a Hammerstein system. The NSR is 10%. Our algorithms with n = 3, m = 1, and d = 2 yield fast convergence and high accuracy as depicted in Fig. 2. This plot shows that our algorithms converge quickly to rather accurate values of the model parameters in five iterations. The final parameter estimates of the Hammerstein model are a_3 = 2.0564, a_2 = 5.0415, a_1 = 4.0188, $b_1=-0.9878$, c_1 = 1.0028, $c_2=1.5197$, and $c_3=-2.9357$. For a comparison, we employ the iterative approach to identify a discrete version of the Hammerstein model. For this purpose, the same input-output data are sampled at $T=0.25$ and the exact orders and delay are assumed to perform the least-squares parameter estimates. For discrete identification, the ordinary least-squares algorithms at each iteration require merely data at individual time instants. Consequently, the algorithms converge to inaccurate parameter estimates due to the presence of measurement noise as indicated in Fig. 2.

We further investigate the robustness of the proposed algorithms with respect to model structure mismatch, i.e. incorrect orders or delay. Assuming a reduced order of n = 2, our algorithms still lead to very good model predictions as revealed in Fig. 3. Note that an input signal used for model validation is different from the test input.

Example 2

$$3x^{(2)}(t)+4x^{(1)}(t)+x(t)=0.5u(t-1.5)$$

$$y(t)=x(t)-0.5[x(t)]^2+2[x(t)]^3$$

This example is a Wiener system. The NSR is 5%. Our algorithms with n = 2, m = 1, and d = 1.5 yield fast convergence and high accuracy as depicted in Fig. 4. This plot shows that our algorithms converge quickly to rather accurate values of the model parameters in ten iterations. The final parameter estimates of the Wiener model are $a_2=2.9649$, $a_1=3.9991$, b_1 = -0.0039, $b_0=0.4998$, $c_2=-0.4727$, and $c_3=2.1814$. On the contrary, the discrete identification without using the integral transform results in slow convergence and poor parameter estimates as indicated in Fig. 4.

Figure 5 verifies the robustness of the proposed algorithms against incorrect delay. With the wrong setting of d = 0, the algorithms still give satisfactory model predictions.

Example 3 is a nonlinear CSTR system discussed by Henson and Seborg [8] and Lee and Huang [4]. This system is neither Hammerstein type nor Wiener type. However, the proposed algorithms can arrive at a second-order Hammerstein model with $a_2=0.0724$,

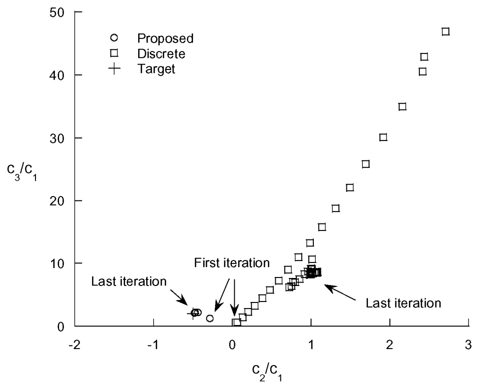


Figure 4. Comparison of the coefficients of the polynomial estimated in the iterative procedures of example 2.

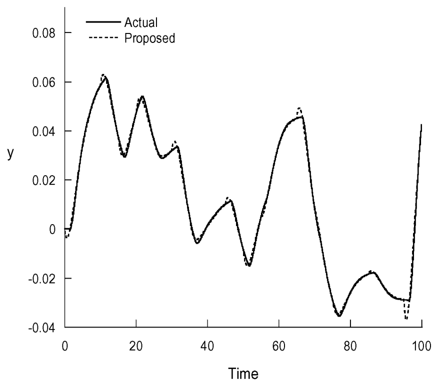


Figure 5. Comparison of the actual response of example 2 with the model predictions obtained by the proposed algorithms with incorrect delay.

$a_1 = 0.2958$, $b_0 = 0.0028$, $c_2 = 5.8589 \times 10^{-5}$, and $c_3 = -7.7572 \times 10^{-8}$. The validity of the identified model is demonstrated by the close agreement between the actual response and the model predictions as seen in Fig. 6. On the other hand, the linear model shows a poor fit to the actual response.

8. Conclusions

It has been demonstrated that the proposed iterative algorithms work well for a variety of process dynamics and test conditions. The algorithms possess satisfactory convergence and accuracy by selecting sufficiently large estimation horizon λ. The use of the integral transform also intensifies the robustness with respect to

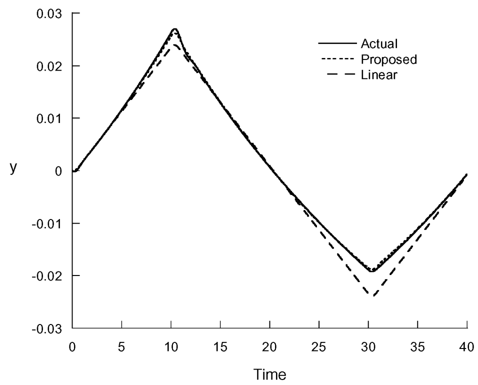


Figure 6. Comparison of the actual response with model predictions for example 3.

measurement noise and model structure mismatch.

Acknowledgement — This work was supported by the National Science Council of R.O.C. under Grant NSC-93-2214-E006-017.

References

[1] Murray-Smith, R. and T. A. Johansen, *Multiple Model Approaches to Modeling and Control*, Taylor and Francis, Landon, 1997.

[2] Haber, R. and H. Unbehauen (1990), "Structure Identification of Nonlinear Dynamic Systems－A Survey on Input/Output Approaches," *Automatica*, **26**, 651.

[3] Huang, H. P., M. W. Lee, and Y. T. Tang (1998), "Identification of Wiener Model Using Relay Feedback Test," *J. Chem. Eng. Jpn*., **31**, 604.

[4] Lee, M. W. and H. P. Huang (1998), "Identification and Control of Hammerstein-Type Nonlinear Process," *J. Chin. Inst. Chem. Engrs.,* **31**, 604.

[5] Voros, J. (1999), "Iterative Algorithm for Parameter Identification of Hammerstein System with Two-Segment Nonlinearities," *IEEE Trans. Automat. Contr*., **44**, 2145.

[6] Voros, J. (2003), "Modeling and Identification of Wiener Systems with Two-Segment Nonlinearities," *IEEE Trans. Contr. Sys. Tech*., **11**, 253.

[7] Hwang, S. H. and M. L. Lin (2003), "Unbiased Identification of Continuous Time Parametric Models Using a Time-Weighted Integral Transform," *Chem. Eng. Comm.,* **190**, 1170.

[8] Henson, M. A. and D. E. Seborg (1990), "Input-Output Linearization of General Nonlinear Processes," *AIChE J*., **36**, 1753.