THE IDENTIFICATION OF A CLASS OF NONLINEAR SYSTEMS USING A CORRELATION ANALYSIS APPROACH

Z Q Lang, M Futterer, and S A Billings

Department of Automatic Control and Systems Engineering
The University of Sheffield, Mappin Street
Sheffield, S1 3JD, UK

Abstract: In this paper, a new correlation analysis based algorithm is proposed for the identification of a class of nonlinear systems which can be described by the NARX (Nonlinear AutoRegressive with eXogeneous input) model with input nonlinearities. Without any assumptions about the structure of an approximating function for the system nonlineariy, the algorithm recovers the functional values of the nonlinearity over a discrete point set associated with the levels of the applied input and estimates the model parameters from the system input output data. An optimal approximating polynomial can then be determined from the nonlinear functional values to produce an optimal estimate for the system nonlinearity. Simulation studies are included to demonstrate the effectiveness of the new method.

Keywords: Identification Algorithm, Nonlinear Models, Nonparametric Identification.

1. INTROCTION

The NARX (Nonlinear AutoRegressive with eXogeneous input) model, which is a relatively general description for dynamic nonlinear systems, was proposed by Billings and co-workers (Billings and Chen, 1989; Chen and Billings, 1989) and has been used to describe the dynamic behaviours of a wide range of engineering systems and physical processes (Billings and Coca, 2002). In this study NARX models which are restricted to nonlinear effects in the input only will be considered. A considerable range of practical systems can be represented by this class of model. The well-known Hammerstein model of nonlinear systems, which is composed of a static nonlinearity followed by a linear dynamic element, is a specific case; The Volterra series can also be represented by this model (Kotsios, 1997).

Because of the complexity of nonlinear systems, the identification of these systems is often addressed based on a specific model structure. For example, the study of the identification of block oriented nonlinear systems such as the Hammerstein and Wiener models. The Wiener model is a cascade system with a linear dynamic element followed by a static nonlinearity, which is just the converse of the Hammerstein model. However, in contrast to the many results for the Hammerstein and Weiner models, as far as we are aware, there is no methods which have been introduced specifically for the identification of the NARX model with input nonlinearities. Although the identification approaches developed for the general NARX model can be used to address this identification problem, restricting the nonlinearities to be in the input only offers several advantages.

In this paper, a new algorithm is proposed for the identification of the NARX model restricted to nonlinearities in the input terms only. The algorithm is a correlation analysis based approach and is derived by extending the idea proposed by the author for the identification of the Hammerstein model (Lang, 1990, 1993, 1994, 1997) to the more general case. Unlike the identification methods for the general NARX model, the new algorithm does not require a priori assumptions for the structure of the approximating function for the system nonlinearity. The algorithm first recovers the functional values of the nonlinearity over a discrete point set associated with the levels of the applied input, and estimates the model parameters directly from the system input output data. Then an optimal approximating polynomial is determined from the obtained nonlinear functional values to produce an optimal estimate for the system nonlinearity. A simulation study is conducted to demonstrate the effectiveness of the proposed method. Theoretical analysis of the properties of the new algorithm is under study, the results will be presented in a future publication.

2. NARX MODEL WITH INPUT NONLINEARITIES

The general NARX model is given by

$$y(k) = f\big(y(k-1),..., y(k-n), u(k-1),..., u(k-m-1)\big) + \varepsilon(k) \qquad (1)$$

and was proposed in 1980's to represent a wide class of discrete time nonlinear systems (Billings and Chen, 1989; Chen and Billings, 1989). In (1), k denotes discrete time, $y(k)$ and $u(k)$ are the system output and input, n and m are the maximum lags for the output and input in the model, $\varepsilon(k)$ represents an i.i.d. sequence, and $f(.)$ is a nonlinear function of its arguments.

The NARX model with input nonlinearities can be described as

$$y(k) = -a_1 y(k-1) -,...,- a_n y(k-n) + q^{-1} g\big(u(k),..., u(k-m)\big) + \varepsilon(k) \qquad (2)$$

where $g(.)$ is a nonlinear function of $u(k-i)$, i=0,…,m. Obviously model (2) is a special case of model (1), but it can still represent a wide range of nonlinear systems including the well-known Hammerstein model and the widely applied Volterra series, as special cases (Kotsios, 1997).

The identification of model (2) involves determining estimates for $a_1,\cdots,a_n$ and the nonlinear function g(.) from the system input and output. This can be carried out using the well established approaches for the general NARX model (Billings and Coca 2002), which expand function g(.) as a polynomial function of its arguments, and estimate $a_1,\cdots,a_n$ and the coefficients of the approximating polynomial via a LS based method. In the next section an alternative approach, based on a new correlation analysis algorithm, is developed specially for the identification of model (2). Without any assumption for the form or structure of the function g(.), the algorithm directly determines the estimates of parameters $a_1,\cdots,a_n$ and the values of the function g() over an discrete point set on the (m+1)-dimensional space of the delayed system inputs $\{u(k),\cdots,u(k-m)\}$. Then an optimal approximating polynomial for the function g(.) is obtained from the functional values. The advantage of the new algorithm is that no indirect modelling errors are involved when estimating the model parameters and functional values. Consequently, estimates of $a_1,\cdots,a_n$ and of an approximating polynomial for g(.) can be obtained with theoretically guaranteed good properties.

3. NEW IDENTIFICATION ALGORITHM

3.1 Derivation of some basic relationships

Consider the NARX model with input nonlinearities given by equation (2). Assume the maximum lags n and m for the system output and input are all known a priori, $\{u(k)\}$ is an i.i.d. random process, and $\{\varepsilon(k)\}$ is a zero mean noise sequence which is independent from the system input $\{u(k)\}$.

Rewrite (2) as

$$y(k+\overline{m}) = -a_1 y(k+\overline{m}-1) -,...,- a_n y(k+\overline{m}-n) + x(k+\overline{m}-1) + \varepsilon(k+\overline{m}) \qquad (3)$$

where $x(k) = g\big(u(k),\cdots,u(k-m)\big)$, and multiply by a manipulated input $G(u(k))$ on both sides of (3) to yield

$$\begin{aligned} y(k+\overline{m})G(u(k)) &= -a_1 y(k+\overline{m}-1)G(u(k)) \\ &-,...,- a_n y(k+\overline{m}-n)G(u(k)) \\ &+ x(k+\overline{m}-1)G(u(k)) + \varepsilon(k+\overline{m})G(u(k)) \end{aligned} \qquad (4)$$

where G(.) is a known function of u(k) such that $E\{G(u(k))\} = 0$. Taking the mathematical expectation on both sides of (4) yields

$$\phi_{yG(u)}(\overline{m}) = -a_1 \phi_{yG(u)}(\overline{m}-1) -,...,- a_n \phi_{yG(u)}(\overline{m}-n) + \phi_{xG(u)}(\overline{m}-1) \qquad (5)$$

where

$$\phi_{yG(u)}(\overline{m}-i) = E\{y(k+\overline{m}-i)G(u(k))\}$$
$$\phi_{xG(u)}(\overline{m}-1) = E\{x(k+\overline{m}-1)G(u(k))\}$$

Consider

$$\phi_{xG(u)}(i) = \begin{cases} \phi_{xG(u)}(i) & \text{when } 0 \le \mathrm{i} \le \mathrm{m} \\ 0 & \text{when } \mathrm{i} > \mathrm{m} \end{cases} \qquad (6)$$

Taking $\overline{m}$ in (5) as 2,3,…,$\overline{M}$ with $\overline{M} > n+m+1$ yields $\overline{M}-1$ equations which can be written in a matrix form as

$$\psi = \Phi\theta \qquad (7)$$

where

$$\psi = \left[\phi_{yG(u)}(2),\cdots,\phi_{yG(u)}(\overline{M})\right]^T$$

$$\Phi = \left[\begin{bmatrix} -\phi_{yG(u)}(1) & \cdots & -\phi_{yG(u)}(2-n) \\ \vdots & \vdots & \vdots \\ -\phi_{yG(u)}(\overline{M}-1) & \cdots & -\phi_{yG(u)}(\overline{M}-n) \end{bmatrix} \begin{bmatrix} [I_{m\times m}] \\ \begin{bmatrix} O \\ \vdots \\ O \end{bmatrix} \end{bmatrix} \right]$$

with $[I_{m\times m}]$ denoting a m dimensional unit matrix and

$$\begin{bmatrix} O \\ \vdots \\ O \end{bmatrix}$$

representing a $(\overline{M}-1-m)\times m$ dimensional zero matrix, and

$$\theta = [a_1, \cdots, a_n, \phi_{xG(u)}(1), \cdots, \phi_{xG(u)}(m)]^T$$

From (7) it is known that the model parameters $a_1, \cdots, a_n$ can be expressed in terms of Φ and ψ as

$$[a_1, \cdots, a_n]^T = [I_{n\times n}\ O_{n\times m}]\ \{\Phi^T \Phi\}^{-1} \Phi^T \psi \tag{8}$$

Equation (8) is the basic relationship based on which the estimates of parameters $a_1, \cdots, a_n$ can be determined using the system input output data.

In order to derive a basic relationship which can be used to determine the functional values of g(.) from the system input output data, rewrite (2) as

$$y(k) = \frac{q^{-1}}{A(q^{-1})} x(k) + \eta(k) = \sum_{i=1}^{\infty} h_i x(k-i) + \eta(k) \tag{9}$$

where

$$A(q^{-1}) = 1 + a_1 q^{-1} +, \cdots, + a_n q^{-n}$$
$$\eta(k) = \varepsilon(k) / A(q^{-1})$$

and $h_i, i = 1, \ldots, \infty$ is the impulse response of the discrete time system transfer function $q^{-1}/A(q^{-1})$.

Replacing k in (9) by $k + \overline{m}$ and taking the conditional expectation with respect to

$$\underline{U}(k) = \{u(k), \cdots, u(k-m)\}$$

on both sides of the equation gives

$$E\{y(k+\overline{m} | \underline{U}(k)\} = \sum_{i=1}^{\infty} h_i E\{x(k+\overline{m}-i) | \underline{U}(k)\} \tag{10}$$

Taking $\overline{m}$ in (10) as

$$(1-m), 1-(m-1), \cdots, 1, \cdots, (1+m)$$

respectively yields $2m+1$ equations. After some manipulations, these equations can be written in a matrix form as below.

$$E_{y\underline{U}} = H_1 E_{x\underline{U}} + H_2 E\{g(\underline{U}(k)\} \tag{11}$$

where

$$E_{y\underline{U}} = \begin{bmatrix} E\{y(k+1-m) | \underline{U}(k)\} \\ \vdots \\ E\{y(k+1) | \underline{U}(k)\} \\ \vdots \\ E\{y(k+1+m) | \underline{U}(k)\} \end{bmatrix} \tag{12}$$

$$E_{x\underline{U}} = \begin{bmatrix} E\{x(k-m) | \underline{U}(k)\} \\ \vdots \\ E\{x(k) | \underline{U}(k)\} \\ \vdots \\ E\{x(k+m) | \underline{U}(k)\} \end{bmatrix} \tag{13}$$

$$H_1 = \begin{bmatrix} h_1 & 0 & \cdots & 0 & \cdots & 0 \\ h_2 & h_1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ h_{m+1} & h_m & \cdots & h_1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \ddots & \vdots \\ h_{2m+1} & h_{2m} & \cdots & h_{m+1} & \cdots & h_1 \end{bmatrix} \tag{14}$$

$$H_2 = \begin{bmatrix} h - h_1 \\ h - h_1 - h_2 \\ \vdots \\ h - \sum_{i=1}^{m+1} h_i \\ \vdots \\ h - \sum_{i=1}^{2m+1} h_i \end{bmatrix} \tag{15}$$

From (11)

$$E_{x\underline{U}} = H_1^{-1} \{E_{y\underline{U}} - H_2 E\{g(\underline{U}(k)\}\} \tag{16}$$

and from (16), $E\{x(k) | \underline{U}(k)\}$, the (m+1)th element of vector $E_{x\underline{U}}$ can be expressed as

$$E\{x(k) | \underline{U}(k)\} = [0_{m\times 1}, 1, 0_{m\times 1}] H_1^{-1} \{E_{y\underline{U}} - H_2 E\{g(\underline{U}(k)\}\} \tag{17}$$

Given a specific

$$U = \{u_1, \cdots, u_{m+1}\},$$

It is straightforward that

$$E\{x(k) | \underline{U}(k) = U\} = E\{g(\underline{U}(k)) | \underline{U}(k) = U\} = g(U) \tag{18}$$

so that (17) gives the value of the function g(.) at the point $U = \{u_{1,\cdots,} u_{m+1}\}$ in the (m+1)-dimensional space of the delayed system inputs $\{u(k), \cdots, u(k-m)\}$. Therefore equation (17) provides the basic relationship needed to determine estimates of the values of the function g().

3.2 Identification algorithm

From equations (8) and (17), a new algorithm for the identification of the NARX model with input nonlinearity given by equation (2) is proposed as follows.

(1) Apply a sequence of i.i.d. random variables with probability distribution

$$u(k) = \begin{cases} D_1 & w.p.1/N \\ D_2 & w.p.1/N \\ \vdots & \\ D_N & w.p.1/N \end{cases} \tag{19}$$

as $\{u(k)\}$ in the experiment performed for the identification. D_i, $i = 1, \cdots, N$ are calculated from the formula

$$D_i = \left\{\cos\frac{(2i-1)\pi}{2N} - \frac{a+b}{a-b}\right\}\frac{b-a}{2} \quad \text{i=1,…,N} \tag{20}$$

where $[a, b]$ is the interval where the system input varies in practice.

This specific choice of input sequence follows the idea in (Lang, 1997). The objective is to determine a best Tchebycheff approximating polynomial for the function g(.) from the values of this function over the discrete point set

$$\Sigma = \left\{ D_{i_1}, D_{i_2}, \cdots\cdots D_{i_{m+1}} \;\middle|\; \begin{matrix} D_{i_j} \in \{D_1, \cdots, D_N\} \\ j = 1, \cdots, m+1 \end{matrix} \right\}$$

on the (m+1)-dimensional space of the delayed system inputs $\{u(k), \cdots, u(k-m)\}$.

(2) From (8), calculate the estimates of parameters $a_1, \cdots, a_n$ as

$$[\hat{a}_1, \cdots, \hat{a}_n]^T = [I_{n\times n}\ O_{n\times m}]\ \{\hat{\Phi}^T\hat{\Phi}\}^{-1}\hat{\Phi}^T\hat{\psi} \tag{21}$$

where $\hat{\Phi}$ and $\hat{\psi}$ are the estimates of Φ and ψ. $\hat{\Phi}$ and $\hat{\psi}$ can be obtained by replacing $\phi_{yG(u)}(\tau)$ in these matrices with the estimate

$$\hat{\phi}_{yG(u)}(\tau) = \frac{1}{(\overline{N}-\tau)}\sum_{k=1}^{\overline{N}-\tau} y(k+\tau)G(u(k)) \tag{22}$$

where $\overline{N}$ is the number of observations for the identification.

(3) From (17), evaluate the estimates of the values of g(.) over the discrete point set Σ as below

$$\hat{g}(\underline{D}) = [0_{m\times 1}, 1, 0_{m\times 1}]\hat{H}_1^{-1}\left\{\begin{matrix} \hat{E}_{y\underline{U}}\big|_{\underline{U}(k)=\underline{D}} \\ -\hat{H}_2\hat{E}\{g(\underline{U}(k)\} \end{matrix}\right\} \tag{23}$$

where $\underline{D} \in \Sigma$, $\hat{H}_1$ and $\hat{H}_2$ are the estimates of H_1 and H_2, and can be obtained directly from the estimates of $a_1, \cdots, a_n$,

$$\hat{E}\{g(\underline{U}(k)\} = \frac{\hat{E}\{y(k)\}}{\sum_{i=1}^{\infty}\hat{h}_i} = (1 + \hat{a}_1 + \cdots + \hat{a}_n)\frac{\sum_{k=1}^{\overline{N}} y(k)}{\overline{N}} \tag{24}$$

and

$$\hat{E}_{y\underline{U}}\big|_{\underline{U}(k)=\underline{D}} = \begin{bmatrix} \hat{E}\{y(k+1-m)|\underline{U}(k)=\underline{D}\} \\ \vdots \\ \hat{E}\{y(k+1)|\underline{U}(k)=\underline{D}\} \\ \vdots \\ \hat{E}\{y(k+1+m)|\underline{U}(k)=\underline{D}\} \end{bmatrix} \tag{25}$$

(4) Determine a H degree best Tchebycheff approximating polynomial $P_H(\underline{U}(k), \theta)$ for the function $g(\underline{U}(k))$ by solving the following optimisation problem

$$\min_{\{\theta\}}\max_{\underline{D}\in\Sigma}\left|\hat{g}(\underline{D}) - P_H(\underline{D}, \theta)\right| \tag{26}$$

where θ is a vector composed of the coefficients of the approximating polynomial, and $H < N$. Denote the solution to (26) as θ^*. Then $P_H(\underline{U}(k), \theta^*)$ is the estimate of the nonlinear function g(.) determined by the algorithm .

This algorithm is a natural extension of the identification method for the Hammerstein model proposed by the author (Lang, 1997) to the NARX model with input nonlinearities. Following the same idea of theoretical analysis in the previous paper, the convergence issues of the new algorithm are under investigation. The results will be presented in a later publication.

4. SIMULATION STUDY

Consider a specific NARX model with input nonlinearities such that

$$y(k) = \frac{q^{-1}}{A(q^{-1})} g(u(k), u(k-1)) + \eta(k)$$

where $A(q^{-1}) = 1 - 0.8q^{-1}$,

$$g(u(k), u(k-1)) = (1.2|u(k)| + 0.24|u(k-1)|)u(k-1)$$

and $\{\eta(k)\}$ is a noise with zero mean and standard deviation 0.01.

Assume that [a,b]=[-1,1], and take N=10, H=8, and $\overline{N} = 2000$. Apply the proposed algorithm to the identification of this system with the function G(u(k)) chosen as

$$G(u(k)) = u(k) + |u(k)| - E\{u(k) + |u(k)|\}$$

The estimation result for $A(q^{-1})$ was

$$\hat{A}(q^{-1}) = 1 - 0.78\,q^{-1}$$

which, clearly , is very close to the true result.

An 8th order approximating polynomial for the nonlinear function g(u(k),u(k-1)) was determined. Figure 1 shows the approximating polynomial over the region of $u(k) \in [-1,1]$ and $u(k-1) \in [-1,1]$. Figure 2 shows the real g(u(k),u(k-1)) over this region. A comparison of Figure 1 and Figure 2 indicates that a very good estimation result for the nonlinear function has be obtained. Therefore the effectiveness of the new algorithm is verified by the simulation example.

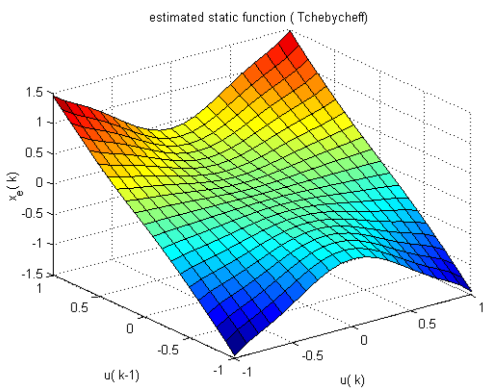

Figure 1 The estimate for the system nonlinearity

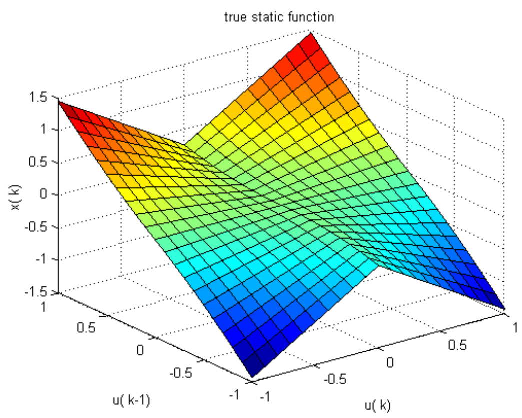

Figure 2 The real system nonlinearity

5. CONCLUSIONS

In the present study, a new algorithm has been developed for the identification of the NARX model with input nonlinearities. The algorithm is a correlation analysis based method and a natural extension of the author's previous work on the identification of the Hammerstein model (Lang, 1997) to a more general case. The simulation study demonstrates the effectiveness of the algorithm. Theoretical analysis of the properties of the algorithm will be discussed in a future publication.

ACKNOWLEDGEMENT

The authors gratefully acknowledge that part of this work was supported by the Engineering and Physical Science Research Council, UK.

REFERENCES

Billings, S.A. and S. Chen (1989). Extended model set, global data and threshold model identification of severely nonlinear systems. *Int. Journal of Control*, **50**, 1897-1923.

Billings S.A. and D. Coca (2002). Identification of narmax and related models. In: *UNESCO Encylopedia of Life Support Systems*, Chapter 10.3, pp 1-11; http://www.eolss.net.

Chen, S. and S. A. Billings (1989). Representations of nonlinear systems: The NARMAX model. *Int. Journal of Control*, **49**, 1013-1032.

Lang, Z. Q. (1997). A nonparametric polynomial identification algorithm for the Hammerstein system. *IEEE Transactions on Automatic Control*, **AC-42**, 1435-1441.

Lang, Z. Q. (1994). On identification of controlled plants described by the Hammerstein system. *IEEE Transactions on Automatic Control,* **AC-39**, 569-573.

Lang, Z. Q. (1993). Controller design oriented model identification method for Hammerstein system. *Automatica*, **29**, pp767-771.

Lang, Z. Q. (1990). A simple method for identification of linear dynamics with hysteresis nonlinear input. *Lecture Notes in Control and Information Sciences*, **144**, 507-515.

Kotsios, S. (1997). Finite input/output representative of a class of Volterra polynomial systems. *Automatica*, **33,** 257-262.

Kotsios, S. and N. Kalouptsidis (1993). The model matching problem for a certain class of nonlinear systems. *Int. Journal of Control*, **57**, 881-919.

Pearson, R. K. (1994). *Discrete time dynamic models*. Oxford University Press.