Systems & Control Letters 83 (2015) 28–37

Contents lists available at ScienceDirect

Systems & Control Letters

journal homepage: www.elsevier.com/locate/sysconle

Fixed point iteration in identifying bilinear models

Guoqi Li [a,*], Changyun Wen [b], Aiming Zhang [c]

[a] Center for Brain Inspired Computing Research, Department of Precision Instrumentation, Tsinghua University, Beijing, 100084, PR China
[b] School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798, Singapore
[c] School of Electrical and Information, Xi'an Jiaotong University, Xi'an, 710049, PR China

ARTICLE INFO

Article history:
Received 5 January 2015
Received in revised form
3 June 2015
Accepted 15 June 2015
Available online 13 July 2015

Keywords:
Bilinear models
Wiener–Hammerstein models
LNL systems
Fixed point
Iterative algorithm
Contraction mapping
Kernel machine
Nonlinear systems identification

ABSTRACT

Inspired by fixed point theory, an iterative algorithm is proposed to identify bilinear models recursively in this paper. It is shown that the resulting iteration is a contraction mapping on a metric space when the number of input–output data points approaches infinity. This ensures the existence and uniqueness of a fixed point of the iterated function sequence and therefore the convergence of the iteration. As an application, one class of block-oriented systems represented by a cascade of a dynamic linear (L), a static nonlinear (N) and a dynamic linear (L) subsystems is illustrated. This gives a solution to the long-standing convergence problem of iteratively identifying LNL (Winer–Hammerstein) models. In addition, we extend the static nonlinear function (N) to a nonparametric model represented by using kernel machine.

1. Introduction

One common model that arises in science and engineering is a class of bilinear models [1], especially in nonlinear system identification [2,3], signal processing and classification [4], machine learning/pattern recognition [5], and many other areas of socioeconomics [6] and biology [7]. For example, one class of block-oriented systems [8] consisting of a dynamic linear (L), a static nonlinear (N) and a dynamic linear (L) subsystems in series can be conveniently formulated as bilinear models. Such an LNL cascade system is called a Wiener–Hammerstein system [9,10] and its identification has been widely studied, see for examples, [10–17].

Due to the wide range applications of bilinear models, there is a strong motivation to develop identification algorithms for such models. Among existing schemes, an iterative algorithm originated in [16] has been extensively used. As pointed out in [17], if the iterative algorithm converges, it converges rapidly and is simple to be implemented. In [18] and [19], the convergence property for iterative algorithm is proved for NL and LN systems when the nonlinear functions can be well parameterized. However, the convergence is generally hard to achieve and unknown in identifying bilinear models. In fact, it was pointed out in [11] and [13] that the convergence problem even for LNL systems with a parametric model [17] representing the N part (the static nonlinear subsystem) has been outstanding for long time and is still unsolved. The main difficulty in obtaining the convergence property is that a block-oriented LNL system contains internal variables, which are generally unmeasurable. It is noted that using nonparametric models [14] to represent nonlinear static functions, which is more efficient and general than parametric models especially when the nonlinear functions are non-smooth or discontinuous, makes the convergence property even more difficult to obtain.

In this paper, we propose an algorithm for the identification of bilinear models iteratively, inspired by fixed point [19,20]. The fixed point of a function is a point that is mapped to itself by the function. In many fields, equilibrium is a fundamental concept that can be described in terms of fixed points and the convergence of a sequence can be analyzed. By exploiting the fixed point theory, it can be proven that the iteration produced by the proposed algorithm is a contraction mapping [21] on a metric space when the number of data points approaches infinity. This guarantees the existence and uniqueness of a fixed point of the iterated function sequence. Therefore the convergence of the iteration is successfully established. As an application, we first show that LNL Wiener–Hammerstein models using a nonparametric model

* Corresponding author.
E-mail addresses: liguoqi@mail.tsinghua.edu.cn (G. Li), ecywen@ntu.edu.sg (C. Wen), zhangam@mail.xjtu.edu.cn (A. Zhang).

http://dx.doi.org/10.1016/j.sysconle.2015.06.008
0167-6911/

named "kernel machine" or "kernel method" [22,23] to represent the N part of the bilinear model, and then apply our proposed algorithm to identify them. This enables the long-standing convergence problem of iteratively identifying LNL Wiener–Hammerstein models to be solved in this paper.

The remaining part of the paper is organized as follows. In Section 2, we introduce bilinear models and fixed point theory as well as the iterative identification method in identifying bilinear models. The representation of LNL Wiener–Hammerstein systems as bilinear models and some relevant analysis are shown in Section 3. Several simulation examples are given in Section 4 to show the performance of the proposed algorithm. Finally, the paper is concluded in Section 5.

2. Bilinear models and fixed point theory

In this section, we first present a common model of bilinear systems. Then an iterative algorithm is proposed to achieve the identification objective with available input–output data points. We show that the estimate $\hat{b}$ of an unknown parameter vector b can be represented as $\hat{b} = \mathcal{F}(\hat{b})$, where function $\mathcal{F}(.)$ is obtained from the iterative algorithm. It is established that $\hat{b} = \mathcal{F}(\hat{b})$ has a unique fixed point which corresponds to the true parameter vector b when the number of data points tends to infinity. We will also prove that the sequence $\{\hat{b}(0), \hat{b}(1), \hat{b}(2), \dots, \hat{b}(k) \dots\}$ generated by the iterated function sequence $\{\hat{b}(0), \mathcal{F}(\hat{b}(0)), \mathcal{F}(\mathcal{F}(\hat{b}(0))), \dots\}$ converges to the fixed point as $k \to \infty$.

2.1. Bilinear models

Usually a linear system described in the following form is considered for its simplicity,

$$y_i = \phi^i d + v_i, \quad i = 1, \dots, N \tag{1}$$

where $\phi^i = [\phi^i_1 \dots \phi^i_r] \in R^{1\times r}$ is a known system vector, $d = [d_1 \dots d_r]' \in R^{r\times 1}$ is an unknown parameter vector, and y_i denotes an observation of the system output with unknown noise v_i. Another common yet more general model in science and engineering is the following bilinear model [5]:

$$\begin{aligned} y_i &= \phi^i d + b'\Psi^i a + v_i \\ &= [\phi^i_1 \dots \phi^i_r] d + b' \begin{bmatrix} \Psi^i_{11} & \dots & \Psi^i_{1L} \\ \vdots & \dots & \vdots \\ \Psi^i_{M1} & \dots & \Psi^i_{ML} \end{bmatrix} a + v_i, \end{aligned} \tag{2}$$

where $b = [b_1 \dots b_M]' \in R^{M\times 1}$, and $a = [a_1 \dots a_L]' \in R^{L\times 1}$ are two vectors of unknown parameters with superscript $'$ denoting the transpose, $\phi^i \in R^{1\times r}$ and $\Psi^i \in R^{M\times L}$ for $i = 1, \dots, N$ are sequences of matrices which describes a bilinear map from the parameter space to the observation space. The model is 'bilinear' because when either a or b is fixed, the relationship between y_i and b or a is linear. Here we note that Ψ^i_{jt} for $j = 1, \dots, M$ and $t = 1, \dots, L$ denoting a component of matrix Ψ^i, is usually related to the available input–output data points.

Denote the observation vector as $Y = [y_1 \dots y_N]'$ and the noise vector as $v = [v_1 \dots v_N]'$. We express the bilinear model in a matrix form of (2) given by $Y = F(a, b, d) + v$ where $F(., .)$ denotes the nonlinearity of the bilinear model, and it can be divided into the following two sub-linear models:

$$\begin{aligned} Y &= \mathcal{G}d + A_a b + v \\ &= \mathcal{G}d + A[I_L \otimes a]b + v, \quad \text{if } a \text{ is known} \end{aligned} \tag{3}$$

$$\begin{aligned} Y &= \mathcal{G}d + A^b a + v \\ &= \mathcal{G}d + A[b \otimes I_M]a + v, \quad \text{if } b \text{ and } d \text{ are known} \end{aligned} \tag{4}$$

where

$$\begin{aligned} \mathcal{G} &= \begin{bmatrix} \phi^i_1 & \dots & \phi^i_r \\ \vdots & \dots & \vdots \\ \phi^i_N & \dots & \phi^N_r \end{bmatrix} \\ A_J &= \begin{bmatrix} \Psi^1_{J1} & \dots & \Psi^1_{JL} \\ \vdots & \dots & \vdots \\ \Psi^N_{J1} & \dots & \Psi^N_{JL} \end{bmatrix}, \quad J = 1, \dots, M \\ A &= [A_1 \dots A_J \dots A_M] \in R^{N\times ML} \\ A_a &\triangleq A[I_L \otimes a] = [A_1 a \dots A_J a \dots A_M a] \\ A^b &\triangleq A[b \otimes I_M] = b_1 A_1 + \dots + b_J A_J + \dots + b_M A_M \end{aligned} \tag{5}$$

where $\otimes$ is the Kronecker product operator [24]. Note that $A_a \in R^{N\times M}$ and $A^b \in R^{N\times L}$ are linearly dependent on a and b, respectively. Here I_M and I_L are conformable identity matrices whose dimensions are the same as the dimension of vectors b and a, i.e., M and L, respectively, for multiplication operation.

Our identification objective is to propose an algorithm to iteratively estimate a, b and d in the general bilinear model of (3) and (4), based on sufficiently large number of input–output data pairs. It will be seen that an LNL nonlinear system can be formulated to the form of above bilinear model in Section 3.

Assumption 1. Components of v are independent identical distributed (i.i.d) variables with zero mean and finite variance $D(v_i) = \sigma^2_v$.

Assumption 2. Matrix $[\mathcal{G}\ A] = [\mathcal{G}\ A_1 \dots A_M]$ is full column rank.

Assumption 3. Either $\|b\|_2$ or $\|a\|_2$ is known and the first nonzero entry of b or a is positive.

Remark 1. Assumption 1 requires the noises to be white. Assumption 2 implies that $\rho_1 I \le \frac{1}{N}[\mathcal{G}\ A][\mathcal{G}\ A]' \le \rho_2 I$ where ρ_1 and ρ_2 are positive numbers. Clearly, this has the same implication as that of the input/output signals being persistently exciting (PE) [25]. Note that if matrices $\mathcal{G}$ and A are constructed based on random input and output signals, Assumption 2 is satisfied provided that the row number of $[\mathcal{G}\ A]$ is not less than its column number. More discussions will be given in Remark 5 when applying the proposed algorithm to LNL systems. Assumption 3 is to guarantee a unique representation of the LNL nonlinear system, as any pair of κa and b/κ for some non-zero κ will give the same input–output data.

2.2. Iterative identification algorithm

Denote the estimates of a, b and d as $\hat{a}$, $\hat{b}$ and $\hat{d}$, respectively. We first obtain $\hat{d}$ without knowing a and b. Then we get $\hat{a}(k)$ and $\hat{b}(k)$ iteratively. Note that (4) can be rewritten as

$$Y = \mathcal{G}d + b_1 A_1 a + \dots + b_M A_M a + v = \mathcal{G}d + A\gamma + v \tag{6}$$

where $\gamma = \begin{bmatrix} b_1 a \\ \vdots \\ b_M a \end{bmatrix}$. Then, estimate $\hat{d}$ is obtained as follows:

$$\hat{d} = (\mathcal{G}'(I_N - A(A'A)^{-1}A')\mathcal{G})^{-1}(\mathcal{G}'(I_N - A(A'A)^{-1}A'))Y \tag{7}$$

where I_N is an identity matrix of dimension N. Later we will show how to derive (7) and establish the consistency of $\hat{d}$ in Theorem 2.4. After $\hat{d}$ is obtained, let $\hat{b}(k)$ be the estimate of b at the kth iteration step. When $\hat{d}$ and $\hat{b}(k)$ become available, determining the estimate $\hat{a}(k)$ of a is to solve a linear equation by substituting them into (3).