Journal of Computational and Applied Mathematics 345 (2019) 135–145

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

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

Model recovery for Hammerstein systems using the hierarchical orthogonal matching pursuit method[☆]

Dongqing Wang[a,c,*], Yaru Yan[a], Yanjun Liu[b], Junhang Ding[a,c]

[a] College of Automation and Electrical Engineering, Qingdao University, Qingdao 266071, China
[b] School of Internet of Things Engineering, Jiangnan University, Wuxi 214122, China
[c] Collaborative Innovation Center for Eco-Textiles of Shandong Province, Qingdao 266071, China

ARTICLE INFO

Article history:
Received 13 October 2016
Received in revised form 21 May 2017

Keywords:
Hierarchical identification principle
Hammerstein system
Orthogonal matching pursuit (OMP)
Compressed sensing (CS)
Parameter estimation

ABSTRACT

Most papers concentrate on the parameter identification of Hammerstein systems with known orders. This paper, motivated by the recent developments in sparse approximations, investigates the combined parameter and order determination of Hammerstein systems. The methodology used relies on greedy schemes—the orthogonal matching pursuit (OMP) algorithm in the compressive sensor (CS) theory. In particular, the first step recasts a bilinear Hammerstein system into two fictitious pseudo-regressive sub-systems which respectively contain the parameters of the nonlinear part or the parameters of the linear part by the hierarchical identification principle. The second step adopts a hierarchical orthogonal matching pursuit (H-OMP) selection procedure to interactively select the parameters and orders of the two sub-systems under the frame of the compressive sensor. Finally, the proposed algorithm is tested on a simulation example.

1. Introduction

Nowadays system modeling and identification are very important for nonlinear systems and complex systems [1–5]. The traditional identification methods for Hammerstein nonlinear systems are popular for the last two decades, including the over-parametrization model based methods [6,7], the iterative/recursive identification methods [8–11], the key term separation principle based identification methods [12–15], the hierarchical identification methods [16–18], and the maximum likelihood estimation methods [19–21], etc. Recently, Li used the Levenberg–Marquardt optimization method to estimate parameters for a Hammerstein output error system [22]; Chen adopted a particle swarm optimization algorithm to estimate unknown parameters for a Hammerstein system [23]. But the above conventional identification methods need process thousands of input and output data, this costs a lot of time in data sampling and parameter estimating.

In the past decade, the compressive sensing method based on the sparsity principle has aroused much attention in signal processing field [24–26], it has an advantage of saving computation in recovering parameters of a system, by collecting only a few data. Generally, the compressive sensing method can be described as to reconstruct a S-sparse vector $\boldsymbol{\Theta} \in \mathbb{R}^n$ from linear measurements in $\boldsymbol{\Phi} \in \mathbb{R}^{m\times n}$ and observations in $\boldsymbol{Y} \in \mathbb{R}^m$ under the form: $\boldsymbol{Y} = \boldsymbol{\Phi}\boldsymbol{\Theta}$. By definition, the number of sparsity, observations, components of the unknown signal are decreased, i.e., $S < m < n$. Accurate reconstruction can be achieved by two types of approaches: the greedy algorithms like the thresholding algorithms [27–29] or the orthogonal matching

☆ This work was supported by the National Natural Science Foundation of China under grant 61573205, 61403217, and the Shandong Provincial Natural Science Foundation of China under grant ZR2015FM017.

* Corresponding author.
E-mail addresses: dqwang64@qdu.edu.cn (D.Q. Wang), yryan92@126.com (Y. Yan), yjl@jiangnan.edu.cn (Y. Liu), dingjunhang@163.com (J. Ding).

https://doi.org/10.1016/j.cam.2018.06.016
0377-0427/

pursuit (OMP) algorithms [30–32], and the basis pursuit (BP) algorithms [33–35]. The greedy OMP algorithm imposes an L_0-norm on the sparse vector $\boldsymbol{\Theta}$, works iteratively by picking up the support columns (atoms) in the measurement matrix in a greedy fashion, and has simple and rapid advantages over the BP method which solves its convex relaxation by the standard L_1-norm technique.

In this paper, a hierarchical orthogonal matching pursuit (H-OMP) algorithm is investigated to simultaneously select the orders and parameters of a Hammerstein system. The explorations lie in three main aspects:

- The first is to recast the system into two fictitious pseudo-regressive sub-systems: one contains the parameters of the nonlinear part and the other contains the parameters of the linear part by the hierarchical identification principle.
- The second is to adopt the H-OMP method to interactively select the parameters and orders of these two sub-systems under the framework of the CS theory.
- Comparing with the existing hierarchical least squares (H-LS) method, the proposed H-OMP algorithm is not necessary to collect a lot of data and invest a lot of power on the parameter identification.

The rest of the paper is organized as follows. Section 2 demonstrates the problem formulation of a Hammerstein system. Section 3 presents the H-OMP identification algorithm to interactively select the orders and parameters of these two sub-systems. Section 4 derives the existing hierarchical least squares (H-LS) algorithm for comparison. Section 5 provides a numerical example for the proposed algorithm. Finally, the concluding remarks are involved in Section 6.

2. The problem formulation

The input nonlinear and output linear functions of a Hammerstein system are expressed as

$$x(t) = f[u(t)] = \sum_{k=1}^{n_c} c_k f_k[u(t)], \tag{1}$$

$$y(t) = B(z)x(t) + v(t), \tag{2}$$

where $B(z) = b_1 z^{-1} + b_2 z^{-2} + \cdots + b_{n_b} z^{-n_b}$, $u(t)$ and $y(t)$ are the system input and output, $x(t)$ is an internal variable, $v(t)$ is stochastic white noise with zero mean, the input nonlinearity f is modeled as a linear combination of basis functions f_i.

Assume that the orders n_b and n_c are unknown, we set sufficient length l as orders of the nonlinear/linear functions ($l > n_b, l > n_c$). From the $y - x$ relationship in (2), we get

$$y(t) = b_1 x(t-1) + b_2 x(t-2) + \cdots + b_{n_b} x(t-n_b) + \cdots + b_l x(t-l) + v(t). \tag{3}$$

Substituting Eq. (1) into $x(t-i)$ into Eq. (3) gives:

$$y(t) = b_1 \sum_{k=1}^{n_c,\ldots,l} c_k f_k[u(t-1)] + \cdots + b_{n_b} \sum_{k=1}^{n_c,\ldots,l} c_k f_k[u(t-n_b)] + \cdots + b_l \sum_{k=1}^{n_c,\ldots,l} c_k f_k[u(t-l)] + v(t).$$

Define the information matrix and the parameter vectors:

$$\boldsymbol{F}(t) = \begin{bmatrix} f_1[u(t-1)], \ldots, f_{n_c}[u(t-1)], \ldots, f_l[u(t-1)] \\ \vdots \\ f_1[u(t-n_b)], \ldots, f_{n_c}[u(t-n_b)], \ldots, f_l[u(t-n_b)] \\ \vdots \\ f_1[u(t-l)], \ldots, f_{n_c}[u(t-l)], \ldots, f_l[u(t-l)] \end{bmatrix} \in \mathbb{R}^{l\times l},$$

$$\boldsymbol{b} = [b_1, b_2, \ldots, b_{n_c}, \underbrace{0, \ldots, 0}_{l-n_b}]^{\mathrm{T}} \in \mathbb{R}^l, \quad \boldsymbol{c} = [c_1, c_2, \ldots, c_{n_c}, \underbrace{0, \ldots, 0}_{l-n_c}]^{\mathrm{T}} \in \mathbb{R}^l,$$

then we have

$$y(t) = \boldsymbol{b}^{\mathrm{T}} \boldsymbol{F}(t) \boldsymbol{c} + v(t). \tag{4}$$

It seems difficult to recast the above obtained bilinear system under the CS framework. In this letter, based on the hierarchical identification principle, we recast the bilinear system into two simple pseudo-regressive sub-systems as follows.

By multiplying $\boldsymbol{F}(t)$ with $\boldsymbol{c}$ ($\boldsymbol{F}(t)\boldsymbol{c} := \boldsymbol{F}_c(t) \in \mathbb{R}^{l\times 1}$), the bilinear model (4) is transformed into a pseudo-regressive **sub-system I** about the parameter vector $\boldsymbol{b}$ of the linear part,

$$\textbf{Sub-system I}: \quad y(t) = \boldsymbol{F}_c^{\mathrm{T}}(t)\boldsymbol{b} + v(t). \tag{5}$$

Similarly, by multiplying $\boldsymbol{b}^{\mathrm{T}}$ with $\boldsymbol{F}(t)$ ($\boldsymbol{b}^{\mathrm{T}}\boldsymbol{F}(t) := \boldsymbol{F}_b(t) \in \mathbb{R}^{1\times l}$), the bilinear model (4) is transformed into a pseudo-regressive **sub-system II** about the parameter vector $\boldsymbol{c}$ of the nonlinear part,

$$\textbf{Sub-system II}: \quad y(t) = \boldsymbol{F}_b(t)\boldsymbol{c} + v(t). \tag{6}$$