Automatica 56 (2015) 53–59

Contents lists available at ScienceDirect

Automatica

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

Brief paper

System identification by discrete rational atoms ☆

Qiuhui Chen [a], Weixiong Mai [b], Liming Zhang [b], Wen Mi [c]

[a] Department of Applied Mathematics, Guangdong University of Foreign Studies, Guangzhou, China
[b] Faculty of Science and Technology, University of Macau, Macau, China
[c] School of Mathematical Science, University of Electronic Science and Technology of China, Chengdu, 611731, China

ARTICLE INFO

Article history:
Received 18 June 2014
Received in revised form
24 November 2014
Accepted 8 March 2015
Available online 11 April 2015

Keywords:
Adaptive
Frequency-domain system identification
Discrete rational atoms

ABSTRACT

A novel adaptive frequency-domain system identification method for linear time-invariant systems is proposed in this paper. It finds poles for discrete rational atoms with discrete frequency responses. The theoretical foundation, including adaptive decomposition principle and decomposition convergence rate, is established. The algorithm of the adaptive pursuit is also provided in this paper.

1. Introduction

It is an important practical problem to model dynamical models from measured data in many fields of science. For system identification of stable linear time-invariant (LTI) systems, rational model structures, such as ARX and ARMAX models, are natural choices, because almost all systems can be described by rational transfer functions (Ljung, 1999). However, direct estimation of parameters of ARX, ARMAX model, etc. encounters nonlinear optimal problems, which is not well solved in numerical aspects so far.

In the last decades, there has been great attention in system identification by using rational orthogonal systems, whose general setting in the unit disc $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$ is

$$\mathcal{B}_k(z) = \mathcal{B}_{\{a_1,\ldots,a_k\}}(z) \triangleq \frac{\sqrt{1-|a_k|^2}}{1-\overline{a}_k z}\prod_{l=1}^{k-1}\frac{z-a_l}{1-\overline{a}_l z}, \tag{1}$$

where all a_k $(k = 1, 2, \ldots)$ are in the unit disc $\mathbb{D}$ and $\overline{a}$ means the complex conjugate of a. This system corresponding to the sequence $\{a_k\}$ is dense in the Hardy space $H_p(\mathbb{D})$ if and only if there holds $\sum_{k=1}^{\infty}(1-|a_k|) = \infty$. The system $\{\mathcal{B}_k : k = 1, 2, \ldots\}$ is called a Takenaka-Malmquist (TM) system. It has a long history of development with applications in mathematics and engineering (Bultheel & De Moor, 2000).

By using a TM system, a stable discrete LTI system $G(z)$ is approximated by $\widetilde{G}(z) = \sum_{k=1}^{n}\theta_k\mathcal{B}_k(z)$, where $\{\theta_k\}_{k=1}^{N}$ is an n-tuple of complex numbers to be determined and n is the order of the model structure. There are an ample amount of publications by such approach both in the time and the frequency domains. Different choices of the poles result in different model structures: when all a_k are zero, it reduces to the classical FIR models; when all a_k are identical with a fixed real number, it gives rise to the Laguerre models (Mäkilä, 1990, 1991; Wahlberg, 1999); and when all a_k are equal with a complex number, it relates to the Kautz models (Wahlberg, 1994; Wahlberg & Mäkilä, 1996). There are also studies on the cases where all a_k's are distinct. This case relates with more generalized rational orthogonal basis (GROB) models (Akçay & Ninness, 1998, 1999; Bultheel, Van gucht, & Van Barel, 2010; Heuberger, Van den Hof, & Wahlberg, 2005; Heuberger et al., 2005; Ninness & Gustafsson, 1997; Pintelon & Schoukens, 2012).

Although there are significant advantages by using the TM system approach, finding appropriate selection of the poles in the TM system is a very important issue to be solved. From the optimal approximation point of view, we may not use the poles of the systems as the poles of the TM functions. Recently, the so-called adaptive Fourier decomposition (AFD), a novel strategy of using TM systems, is proposed in Qian (2006); Qian and Wang (2011). The novelty of AFD is that it undergoes a one-by-one selection of the

☆ This work was supported by Macao Government FDCT 098/2012/A3; Multi-Year Research Grant of the University of Macau No. MYRG116(Y1-L3)-FST13-QT; The Natural Science Foundation of Guangdong Province (No. S2011010004986); NSFC grant 11401079. The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Andrea Garulli, under the direction of Editor Torsten Söderström.

E-mail addresses: chenqiuhui@hotmail.com (Q. Chen), maiweixiong@gmail.com (W. Mai), lmzhang@umac.mo (L. Zhang), mathmw@uestc.edu.cn (W. Mi).

http://dx.doi.org/10.1016/j.automatica.2015.03.022
0005-1098/

poles a_k: At each step it performs a maximal selection criterion. It results in a significantly fast rational approximation. Based on AFD, a two-step algorithm for frequency-domain system identification is established in Mi and Qian (2012). In this two-step algorithm, one first constructs a Hardy H^2 function by using a discretized Cauchy integral formula and then produces approximations by AFD. This program is subsequently developed in Mi and Qian (2014) and Mi, Qian, and Wan (2012).

The present paper proposes an algorithm to select poles based on the measured data. In such approach we do not assume there exists a system function. We note that, in AFD, the parameterized and normalized Szegö kernels $e_a(z) = \frac{\sqrt{1-|a|^2}}{1-\bar{a}z}$, $a \in \mathbb{D}$, play a key role. In fact, any infinite orthogonal rational basis $\{\mathcal{B}_j : j = 1, 2, \ldots\}$ can be obtained through the Gram–Schmidt process applied to a set of functions consisting of e_{a_j} and their derivatives.

What is done in this paper is to use samples of normalized parameterized Szegö kernels in the finite sequence form to replace normalized parameterized Szegö kernels as functions. The purpose is to construct an AFD-like algorithm but with the maximal selection criterion in terms of the measured data. The aspect of using Szegö kernels makes it like AFD, while the aspect of being lack of algebraic structure makes it more like matching pursuit. The proposed algorithm may be considered as discretization of AFD, as if DFT in the Fourier series context (see Mi & Qian, 2012; Qian & Wang, 2011). In Ward and Partington (1995, 1996), the authors studied the discrete atoms corresponding to Laguerre models. The collocation matrix corresponding to Laguerre models is nonsingular, since it is a special case of Vandermonde matrix. In the present paper, we will show that the collocation matrix corresponding to a class of Szegö kernels is nonsingular, so is the one corresponding to a class of Blaschke product.

Our presented algorithm contains two steps: choosing poles in discrete system and calculating the weight. About calculation of the weight, there are two approaches to realize: greedy algorithm and least square (LS) method. Throughout the paper, the term *proposed method with greedy weight* (*PMGW*) means that we use the maximal selection criterion to choose poles and use greedy algorithm to calculate the weight. Similarly, the term *proposed method with LS weight* (*PMLSW*) stands for the algorithm combining the maximal selection criterion to choose poles with using LS method to get the weight.

In numerical implementation, we mainly compare the two proposed methods with the Laguerre models and the subspace method. The results show that the two proposed methods perform very well for the example studied here. Particularly, *PMLSW* is better than the other methods, while frequency-domain measurements are corrupted by noise.

The writing plan is as follows. Section 2 is the problem setting. Section 3 is devoted to construction of the discrete Szegö system as well as exploration of its properties. We prove that the collocation matrix is nonsingular and the system forms a frame (even Riesz basis) of the vector space $\mathbb{C}^N$. Section 4 contributes to designing and analyzing the identification algorithm for the adaptive representation. Since the discrete Szegö system contains unknown parameters, the benefit is the flexibility and adaptivity for the presentation of practical data. The main algorithm is based on the maximum selection criterion in greedy algorithm with minor modification. Section 5 is focused on *PMGW* and *PMLSW*. In Section 6, some numerical results are given. The last section draws conclusions.

2. Problem setting

In this paper it is assumed that a set of frequency-domain measurements $\mathbf{y} = \{y_k\}_{k=1}^N$ is available. These measurements are obtained from a single input, single output (SISO) discrete stable LTI system $f(z)$ in $H_2(D)$ with real-valued impulse responses. In general, $f(z)$ is a rational function of real-valued coefficients with its poles being outside the closure of the unit disc $\mathbb{D}$.

It is then assumed that the structure of measurements $\mathbf{y}$ is set up as

$$y_k = f(e^{j\omega_k}) + \upsilon_k \quad (k = 1, 2, \ldots, N), \tag{2}$$

where the samples $\{\omega_k : k = 1, \ldots, N\}$ can be any distinct points in the interval $[0, 2\pi)$. In the equidistant case, it takes $\omega_k = \frac{2\pi(k-1)}{N}$. The positive integer N is assumed to be even. In this paper, the noise (error sequence) $\{\upsilon_k\}$ is a stochastic model with zero mean and variance $\sigma^2 < \infty$.

Let $a_1, \ldots, a_n$ be distinguished points in $\mathbb{D}$, $X_n = span\{\mathcal{B}_1, \ldots, \mathcal{B}_n\} = span\{e_{a_1}(z), \ldots, e_{a_n}(z)\}$, where the orthonormal system $\{\mathcal{B}_k\}_{k=1}^n$ is defined by (1). The identification problem is as follows.

Frequency-domain identification problem: Given a set of frequency-domain measurements $\mathbf{y} = \{y_k\}_{k=1}^N$ for $G \in H^2$. Find an optimal approximation $G_n(z) = \sum_{k=1}^n \theta_k e_{a_k}(z)$ by determining $\{a_k\}_{k=1}^n \in \mathbb{D}$ and coefficients $\{\theta_k\}_{k=1}^n$.

To give a solution of this problem, we start with reviewing relevant properties of the discrete Szegö system.

3. Discrete system based on Szegö kernel

Set the column vector $\mathbf{E}_a$ in $\mathbb{C}^N$ by

$$\mathbf{E}_a = \left(\frac{\sqrt{1-|a|^2}}{1-\bar{a}e^{i\omega_1}}, \ldots, \frac{\sqrt{1-|a|^2}}{1-\bar{a}e^{i\omega_N}}\right)^T. \tag{3}$$

For a parameter sequence $\{a_j \in \mathbb{D} : j = 1, 2, \ldots\}$ with $a_1 = 0$, we can define, with a bit of abuse of notation, the discrete system $\{\mathbf{E}_k : k \in \mathbb{Z}\}$ and the unit system $\{\mathbf{e}_k : k \in \mathbb{Z}\}$ in $\mathbb{C}^N$, respectively, by

$$\mathbf{E}_1 = (1, \ldots, 1)^T, \qquad \mathbf{E}_k = \mathbf{E}_{a_k}, \quad k \in \mathbb{N} \tag{4}$$

and

$$\mathbf{e}_1 = \frac{1}{\sqrt{N}}(1, \ldots, 1)^T, \qquad \mathbf{e}_k = \frac{\mathbf{E}_{a_k}}{h(a_k; \mathbf{w})}, \quad k \in \mathbb{N}, \tag{5}$$

where the sample vector $\mathbf{w} = (\omega_1, \ldots, \omega_N)$, and, $h(\cdot; \mathbf{w}) : \mathbb{C} \to \mathbb{R}$ is the complex-variable and real-valued function defined by

$$h(a; \mathbf{w}) = \left(\sum_{s=1}^N \frac{1-|a|^2}{|1-\bar{a}e^{i\omega_s}|^2}\right)^{\frac{1}{2}}. \tag{6}$$

Define the collocation matrix B by

$$B = (\mathbf{e}_1, \mathbf{e}_1, \ldots, \mathbf{e}_N). \tag{7}$$

The main theorem of this section is given below.

Theorem 1. *Suppose that samples $\{\omega_j : j = 1, \ldots, N\}$ satisfy the condition $e^{i(\omega_j-\omega_k)} \neq 1$ when $j \neq k, j, k \in \{1, \ldots, N\}$, and $a_j, j = 1, \ldots, N$, are distinct points in $\mathbb{D}$ ($a_1 = 0$). Then, the matrix B defined in (7) is nonsingular.*

We offer an outline of the proof of Theorem 1. Firstly, by making use of the fundamental theorem of algebra, we prove that, for any nonzero vector $\hbar = (h_1, \ldots, h_N) \in \mathbb{C}^N$, there exists at least one point $b \in \mathbb{D}$ such that $|\langle \hbar, \mathbf{E_b}\rangle| > 0$. Secondly, we prove the following fact by contradiction:

$$E = (\mathbf{E}_1, \mathbf{E}_2, \ldots, \mathbf{E}_N) \tag{8}$$