J. Math. Anal. Appl. 409 (2014) 494–508

Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

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

Hankel matrices for system identification☆

Bi-Qiang Mu, Han-Fu Chen *

Key Laboratory of Systems and Control, Institute of Systems Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, PR China

ARTICLE INFO

Article history:
Received 30 January 2013
Available online 19 July 2013
Submitted by Xu Zhang

Keywords:
Hankel matrix
Row-full-rank
Impulse response
Correlation function
Multi-variable linear systems

ABSTRACT

The coefficients of a linear system, even if it is a part of a block-oriented nonlinear system, normally satisfy some linear algebraic equations via Hankel matrices composed of impulse responses or correlation functions. In order to determine or to estimate the coefficients of a linear system it is important to require the associated Hankel matrix be of row-full-rank. The paper first discusses the equivalent conditions for identifiability of the system. Then, it is shown that the row-full-rank of the Hankel matrix composed of impulse responses is equivalent to identifiability of the system. Finally, for the row-full-rank of the Hankel matrix composed of correlation functions, the necessary and sufficient conditions are presented, which appear slightly stronger than the identifiability condition. In comparison with existing results, here the minimum phase condition is no longer required for the case where the dimension of the system input and output is the same, though the paper does not make such a dimensional restriction.

1. Introduction

A block-oriented nonlinear system often includes some linear parts as its subsystems, for example, the Hammerstein system is composed of a nonlinear static block followed by a linear subsystem, and the Wiener–Hammerstein system is a nonlinear static block sandwiched by two linear subsystems. When identifying such kind of systems, one has to estimate not only the nonlinearities but also their linear subsystems. From the existing papers, e.g., [3,11,14,8,9] among others, it is seen that the Hankel matrices composed of impulse responses of the linear subsystem as well as composed of the correlation functions of its output are of crucial importance for estimating the unknown coefficients of the system. Let us explain this more clearly.

Consider the following linear model

$$A(z)y_k = B(z)u_k, \tag{1}$$

where

$$A(z) = I + A_1 z + \cdots + A_p z^p \quad \text{with } A_p \neq 0 \tag{2}$$

$$B(z) = B_0 + B_1 z + \cdots + B_q z^q \quad \text{with } B_q \neq 0 \tag{3}$$

are matrix polynomials in the backward-shift operator $z : \; zy_k = y_{k-1}$. The system output y_k and input u_k are of n- and m-dimensions, respectively.

☆ This work was supported by NSFC under Grants 61273193, 61120106011, 61134013, and the National Center for Mathematics and Interdisciplinary Sciences, Chinese Academy of Sciences.

* Corresponding author.
E-mail address: hfchen@iss.ac.cn (H.-F. Chen).

0022-247X/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jmaa.2013.07.025

Assume $A(z)$ is stable, i.e., $\det A(z) \neq 0, \forall |z| \leq 1$.

In time series analysis, it is required to estimate the systems orders (p, q), the matrix coefficients $(A_1, \dots, A_p, B_1 \dots, B_q)$, and the covariance matrix $\Sigma_u = Eu_k u_k^T$ of the innovation process on the basis of output data $\{y_0, y_1, y_2, \dots\}$ for the case where $n = m$, $B_0 = I$, and $\{u_k\}$ is a sequence of zero-mean iid (independent and identically distributed) or uncorrelated random vectors.

In system and control, as a rule $n \geq m$, and it is also required to estimate the systems orders (p, q) and the matrix coefficients $(A_1, \dots, A_p, B_0, B_1 \dots, B_q)$ with B_0 included. If the linear model is a part of nonlinear systems, e.g., the Hammerstein system, the Wiener system, etc., then the available information for identification may be the noisy or estimated inputs and outputs of the model.

Stability of $A(z)$ gives possibility to define the transfer function:

$$H(z) \triangleq A^{-1}(z)B(z) = \sum_{i=0}^{\infty} H_i z^i, \tag{4}$$

where $H_0 = B_0$, $\|H_i\| = O(e^{-ri})$, $r > 0$, $i > 1$. Then, y_k in (1) can be connected with the input $\{u_k\}$ via impulse responses:

$$y_k = \sum_{i=0}^{\infty} H_i u_{k-i}. \tag{5}$$

Let us first derive the linear equations connecting $\{A_1, \dots, A_p, B_0, B_1, \dots, B_q\}$ with $\{H_i\}$.

From (4), it follows that

$$B_0 + B_1 z + \cdots + B_q z^q = (I + A_1 z + \cdots + A_p z^p)(H_0 + H_1 z + \cdots + H_i z^i + \cdots). \tag{6}$$

Identifying coefficients for the same degrees of z at both sides implies

$$B_i = \sum_{j=0}^{i \wedge p} A_j H_{i-j} \quad \forall 0 \leq i \leq q, \tag{7}$$

$$H_i = -\sum_{j=1}^{i \wedge p} A_j H_{i-j} \quad \forall i \geq q+1, \tag{8}$$

where $A_0 = I$ and $a \wedge b$ denotes $\min(a, b)$.

For H_i, $q+1 \leq i \leq q+np$, by (8) we obtain the following linear algebraic equation

$$[A_1, A_2, \dots, A_p]L = -[H_{q+1}, H_{q+2}, \dots, H_{q+np}], \tag{9}$$

where

$$L \triangleq \begin{pmatrix} H_q & H_{q+1} & \cdots & H_{q+np-1} \\ H_{q-1} & H_q & \cdots & H_{q+np-2} \\ \vdots & \vdots & \ddots & \vdots \\ H_{q-p+1} & H_{q-p+2} & \cdots & H_{q+(n-1)p} \end{pmatrix}, \tag{10}$$

where $H_i \triangleq 0$ for $i < 0$.

Define

$$\theta_A^T \triangleq [A_1, \dots, A_p], \qquad W^T \triangleq -[H_{q+1}, H_{q+2}, \dots, H_{q+np}]. \tag{11}$$

Then, from (9) it follows that

$$\theta_A = (LL^T)^{-1} LW, \tag{12}$$

if L is of row-full-rank.

In this case, if we can obtain estimates for $\{H_i\}$, then replacing H_i, $i = 0, 1, 2, \dots$ in (9) with their estimates, we derive the estimate for θ_A. Finally, with the help of (7) the estimates for B_i, $i = 0, 1, \dots, q$ can also be obtained.

From here we see that the row-full-rank of the Hankel matrix L composed of impulse responses is important for estimating the system indeed.

The well-known Yule–Walker equation connects θ_A with the Hankel matrix composed of correlation functions of the system output $\{y_k\}$.

Under the stability assumption on $A(z)$, $\{y_k\}$ is a stationary process with correlation function $R_i \triangleq Ey_k y_{k-i}^T$, if $\{u_k\}$ is a sequence of zero-mean uncorrelated random vectors with the same second moment.

Multiplying y_{k-t}^T, $t \geq q+1$ on the both sides of (1) from right and taking expectation, we obtain

$$E(y_k + A_1 y_{k-1} + \cdots + A_p y_{k-p}) y_{k-t}^T = E(B_0 u_k + B_1 u_{k-1} + \cdots + B_q u_{k-q}) y_{k-t}^T = 0 \quad \forall t \geq q+1,$$