九州大学学術情報リポジトリ
Kyushu University Institutional Repository

Parameter Estimation of Orthonormal Functions Using Block Toeplitz Construction

Bae, Chul-Min
Department of Electrical and Electronic Systems Engineering, Graduate School of Information Science and Electrical Engineering, Kyushu University : Visiting Researcher

Wada, Kiyoshi
Department of Electrical and Electronic Systems Engineering, Faculty of Information Science and Electrical Engineering, Kyushu University

Yang, Zi Jiang
Department of Electrical and Electronic Systems Engineering, Faculty of Information Science and Electrical Engineering, Kyushu University

Jin, Chun-Zhi
Department of Electrical and Electronic Systems Engineering, Faculty of Information Science and Electrical Engineering, Kyushu University

https://doi.org/10.15017/1524317

出版情報：九州大学大学院システム情報科学紀要. 7 (1), pp.7-12, 2002-03-26. 九州大学大学院システム情報科学研究院
バージョン：
権利関係：

九州大学大学院
システム情報科学紀要
第7巻 第1号 平成14年3月

Research Reports on Information Science and
Electrical Engineering of Kyushu University
Vol.7, No.1, March 2002

Parameter Estimation of Orthonormal Functions Using Block Toeplitz Construction

Chul-Min BAE* , Kiyoshi WADA** , Zi Jiang YANG** and Chun-Zhi JIN**

(Received December 14, 2001)

Abstract: In many areas of signal, system, and control theory, orthogonal functions play an important role in issues of analysis and design. In this paper, we will expand and generalize the orthogonal functions as basis functions for dynamical system representations. The orthogonal functions can be considered as generalizations of, for example, the pulse functions, Laguerre functions, and Kautz functions. A least-squares identification method is studied that estimates a finite number of expansion coefficients in the series expansion of a transfer function, where the expansion is in terms of recently introduced generalized orthogonal functions. The analysis is based on the result that the corresponding linear regression normal equations have a block Toeplitz structure. It is shown how we can exploit a block Toeplitz structure to increase the speed of convergence in a series expansion.

Keywords: Kautz functions, Laguerre functions, Block Toeplitz

1. Introduction

The use of orthogonal functions with the aim of adapting the system and signal representation to the specific properties of the systems and signal has a long history. The main part of this work dates back to the classical work of Lee.[8] There has been interest in developing schemes for estimation of single-input single-output systems using so-called orthonormal basis functions. [3) 4) 11)]
The main applications are in system identification and adaptive signal processing, where the parametrization of models in terms of finite expansion coefficients is attractive because of the linear in the parameters model structure. This allows the use of linear regression estimation techniques to identify the system from observed input and output datas.

A model of a linear stable time-invariant system with additive disturbance is given by:

$$y(t) = G^0(q)u(t) + v(t) \tag{1}$$

where

$$G^0(q) = \sum_{k=1}^{\infty} g_k q^{-k}, \tag{2}$$

where $u(t)$ and $y(t)$ are the input and output signals, respectively. Time shifts are represented by the delay operator $q^{-1}u(t) = u(t-1)$. And $v(t)$ is a unit-variance, zero-mean white noise process. System identification deals with the problem of finding an estimate of $G^0(z)$ from observations of $\{y(t), u(t)\}_{t=1,\cdots,n}$, see. [1)] The identification problem simplifies to a linear regression estimation problem if the model can be represented by

$$G(z) = \sum_{k=1}^{\infty} w_k \Psi_k(z), \tag{3}$$

with $\{w_k\}_{k=1,2,\dots}$ the unknown model parameters. There are a number of research areas that deal with the question of either approximating a given system G with a finite number of coefficients in a series expansion as in (3), or identifying an unknown system in terms of a finite number of expansion coefficients through

$$\widehat{G}(z) = \sum_{k=1}^{n} \widehat{w}_k \Psi_k(z) \tag{4}$$

where the accuracy of the model will be essentially dependent on the choice of basis functions $\Psi_k(z)$. Note that the choice $\Psi_k(z) = z^{-k}$ corresponds to the use of so-called FIR (finite impulse response) models.[1)] In section 2 we present the orthonormal functions, and we formulate an identification of expansion coefficients in section 3. In section 4 we introduce the inversion algorithm of block Toeplitz matrix. Finally, the proposed approximation tech-

* Department of Electrical and Electronic Systems Engineering, Visiting Researcher
** Department of Electrical and Electronic Systems Engineering

nique is illustrated by a simple example.

2. Orthonormal Functions

The FIR model $\{z^{-k}\}_{k=1,2,\cdots}$ is not a suitable operator for short sampling intervals, the reason being that it has a too short memory. By using operators with longer memory, the number of parameters necessary to describe useful approximations can be reduced. We shall start by analyzing the continuous-time case, which gives useful information about what can be expected for short sampling intervals. The problem of orthogonalizing a set of continuous time exponential functions has been elegantly solved in [1]. The key idea is to determine the corresponding Laplace transforms, which have very simple structures. The analogous discrete problem is summarized by the following theorem.
The sequence of functions $\{\Psi_k(z)\}$ is determined as follows:

$$\Psi_{2k-1}(z) = C_1^{(k)}(1 - a_1^{(k)}z)\Gamma^{(k)}(z) \tag{5}$$

$$\Psi_{2k}(z) = C_2^{(k)}(1 - a_2^{(k)}z)\Gamma^{(k)}(z) \tag{6}$$

where

$$\Gamma^{(k)}(z) = \frac{\prod_{j=1}^{k-1}(1-\beta_j z)(1-\beta_j^* z)}{\prod_{j=1}^{k}(z-\beta_j)(z-\beta_j^*)}$$

$$C_1^{(k)} = \sqrt{\frac{(1-\beta_k^2)(1-\beta_k^{*2})(1-\beta_k\beta_k^*)}{(1+(a_1^{(k)})^2)(1+\beta_k\beta_k^*) - 2a_1^{(k)}(\beta_k+\beta_k^*)}}$$

$$C_2^{(k)} = \sqrt{\frac{(1-\beta_k^2)(1-\beta_k^{*2})(1-\beta_k\beta_k^*)}{(1+(a_2^{(k)})^2)(1+\beta_k\beta_k^*) - 2a_2^{(k)}(\beta_k+\beta_k^*)}}$$

$$(1 + a_1^{(k)}a_2^{(k)})(1+\beta_k\beta_k^*) - (a_1^{(k)} + a_2^{(k)})(\beta_k+\beta_k^*) = 0 \tag{7}$$

Here β_k^* are complex numbers such that $|\beta_k| < 1$, and $a_1^{(k)}$, $a_2^{(k)}$ are restricted by the condition (7). The functions $\{\Psi_k(z)\}_{k=1,2,\cdots}$ will be called the discrete Kautz functions.

2.1 Laguerre Function

For $\beta_k = a$, where a is real, we can take

$$a_1^{(k)} = a, \quad a_2^{(k)} = \frac{1}{a}$$

$$C_1^{(k)} = (1-a^2)^{\frac{1}{2}}, \quad C_2^{(k)} = a(1-a^2)^{\frac{1}{2}}$$

Then, Kautz functions $\{\Psi_k(z)\}$ simplify to the real discrete Laguerre function

$$\begin{aligned}\Psi_k(z) = L_k(z,a) &= \frac{\sqrt{1-a^2}}{z-a}\left(\frac{1-az}{z-a}\right)^{k-1} \\ &= L \cdot G_L(z)^{k-1}\end{aligned} \tag{8}$$

where

$$L = \frac{\sqrt{1-a^2}}{z-a}, \qquad G_L(z) = \frac{1-az}{z-a}.$$

System identification using Laguerre models is studied in detail in [3]. By taking $a = 0$, the Laguerre function simplifies to an ordinary finite impulse response (FIR) model.

2.2 Kautz Function

Another special case is for $\beta_k = \beta$. For this case one can take:

$$a_1^{(k)} = \frac{1+\beta\beta^*}{\beta+\beta^*}, \quad a_2^{(k)} = 0$$

and thus

$$\begin{aligned}\Psi_{2k-1}(z) &= \frac{\sqrt{1-c^2}(z-b)}{z^2+b(c-1)z-c} \left[\frac{-cz^2+b(c-1)z+1}{z^2+b(c-1)z-c}\right]^{k-1} \\ &= K_{2k-1}(z)G_k(z)^{k-1}\end{aligned}$$

$$\begin{aligned}\Psi_{2k}(z) &= \frac{\sqrt{(1-c^2)(1-b^2)}}{z^2+b(c-1)z-c} \left[\frac{-cz^2+b(c-1)z+1}{z^2+b(c-1)z-c}\right]^{k-1} \\ &= K_{2k}(z)G_k(z)^{k-1}\end{aligned}$$

where

$$K_{2k-1}(z) = \frac{\sqrt{1-c^2}(z-b)}{z^2+b(c-1)z-c}$$

$$K_{2k}(z) = \frac{\sqrt{(1-c^2)(1-b^2)}}{z^2+b(c-1)z-c}$$

$$G_k(z) = \frac{-cz^2 + b(c-1)z + 1}{z^2 + b(c-1)z - c}$$

and $b = (\beta + \beta^*)/(1 + \beta\beta^*)$, $c = -\beta\beta^*$. Since $a_1^{(k)}$ and $a_2^{(k)}$ are not unique, several other sets of $\{\Psi_k(z)\}$ are possible.

3. Expansion Coefficients

Using Laguerre and Kautz function a practical parameter identification method for linear time-invariant systems is introduced.
A linear model structure will be employed, determined by

$$\widehat{y}(t) = \sum_{k=1}^{N} \widehat{w}_k \Psi_k(z) u(t) \tag{9}$$

Given data $\{u(t), y(t)\}_{t=1,\cdots,N}$ taken from experiments on this system, the corresponding prediction error is given by

$$\varepsilon(t) = y(t) - \sum_{k=1}^{N} \widehat{w}_k \Psi_k(z) u(t) \tag{10}$$

The corresponding estimated transfer function by

$$\widehat{G}(z) = \sum_{k=1}^{N} \widehat{w}_k \Psi_k(z) \tag{11}$$

where $\{\Psi_k(z)\}$ is a set of given basis functions and $\{\widehat{w}_k\}$ are the unknown model parameters. The least squares method can now be applied to estimate the model parameters

$$\theta^T = (\widehat{w}_1, \widehat{w}_2, \cdots, \widehat{w}_N) \quad . \tag{12}$$

The input and output relation can be written in the linear regression form

$$y(t) = z_t^T \theta \tag{13}$$

where

$$z_t^T = [\bar{u}_1(t), \bar{u}_2(t), \cdots, \bar{u}_n(t)]$$

$$\bar{u}_k(t) = \Psi_k(z) u(t),$$

Let

$$Z^T = [z_1, \cdots, z_N], \quad \mathbf{y}^T = [y(1), \cdots, y(N)].$$

Then, the least squares estimate of θ minimizes the loss function is such as:

$$J = \frac{1}{N} \sum_{t=t_0}^{N} (y(t) - z_t^T \theta)^2$$

$$= \frac{1}{N} (\mathbf{y} - Z\theta)^T (\mathbf{y} - Z\theta). \tag{14}$$

The solution of this quadratic optimization problem is:

$$\hat{\theta}_N = (R_N)^{-1} f_N \tag{15}$$

where

$$R_N = \frac{1}{N} \sum_{t=t_0}^{N} z_t z_t^T,$$

$$f_N = \frac{1}{N} \sum_{t=t_0}^{N} z_t y(t).$$

The value of t_0 depends on how the effects of unknown initial conditions are treated. For large N, the effects of t_0 will be negligible. R_N is represented as follows

$$R_N = \begin{bmatrix} U_{11} & U_{12} & \cdots & U_{1n} \\ U_{21} & U_{22} & \cdots & U_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ U_{n1} & U_{n2} & \cdots & U_{nn} \end{bmatrix} \tag{16}$$

where

$$U_{ij} = \begin{bmatrix} \overline{u}_{2i-1}^T \overline{u}_{2j-1} & \overline{u}_{2i-1}^T \overline{u}_{2j} \\ \overline{u}_{2i}^T \overline{u}_{2j-1} & \overline{u}_{2i}^T \overline{u}_{2j} \end{bmatrix}$$

$$(i=1,2,\cdots,j=1,2,\cdots).$$

Where R_N is the block Toeplitz matrix of dimension $n \times n$.[4)]

4. The Inversion Algorithm of Block Toeplitz Matrix

Consider R_N is a block Toeplitz matrix when we estimate a Kautz function. Let the block Toeplitz matrix R_N be composed of $n \times n$ blocks of size $p \times p$ that is,

$$R_{p,n} = \begin{bmatrix} R_0 & R_1 & \cdots & R_{n-1} \\ R_{-1} & R_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & R_1 \\ R_{-n+1} & \cdots & R_{-1} & R_0 \end{bmatrix} \tag{17}$$

where, R_j is the $p \times p$ matrix

$$R_j = \begin{bmatrix} r_0 & r_1 & \cdots & r_{p-1} \\ r_{-1} & r_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & r_1 \\ r_{-p} & \cdots & r_{-1} & r_0 \end{bmatrix}. \quad (18)$$

Let denote by $J_{p,n}$ the exchange matrix, given by

$$J_{p,n} = \begin{bmatrix} \mathbf{0} & & J_p \\ & \cdot^{\cdot^{\cdot}} & \\ J_p & & \mathbf{0} \end{bmatrix} \quad (19)$$

where J_j is the $p \times p$ matrix with ones in the antidiagonal and zeros elsewhere

$$J_j = \begin{bmatrix} \mathbf{0} & & 1 \\ & \cdot^{\cdot^{\cdot}} & \\ 1 & & \mathbf{0} \end{bmatrix}. \quad (20)$$

From (16), and (17) it follows that the sequence of matrices has a nested structure

$$R_{p,n+1} = \begin{bmatrix} R_{p,n} & \underline{R}_n \\ \underline{R}_n^T & R_0 \end{bmatrix} \quad (21)$$

where

$$\underline{R}_n = \begin{bmatrix} R_n \\ \vdots \\ R_1 \end{bmatrix} \qquad \underline{R}_n^T = [R_n^T \cdots R_1^T]. \quad (22)$$

Applying the well-known the matrix inversion lemma for a partitioned matrix to (22), (23), we obtain

$$R_{p,n+1}^{-1} = \begin{bmatrix} R_{p,n}^{-1} + \underline{W}_n \alpha_n^{-1} \underline{V}_n^T & \underline{W}_n \alpha_n^{-1} \\ \alpha_n^{-1} \underline{V}_n^T & \alpha_n^{-1} \end{bmatrix} \quad (23)$$

where

$$\underline{W}_n = -R_{p,n}^{-1} \underline{R}_n \quad (24)$$

$$\underline{V}_n^T = -\underline{R}_n^T R_{p,n}^{-1} \quad (25)$$

and

$$\alpha_n = R_0 - \underline{R}_n^T R_{p,n}^{-1} \underline{R}_n. \quad (26)$$

Thus, a recursive scheme for the inversion of $R_{p,n}$ is possible, provided we can derive recursive schemes for the computation of the matrices W_n,V_n^T and α_n. This aim let us first introduce the following notation.

$$\widehat{R}_n \triangleq J_p R_n, \qquad \widehat{R}_n^T \triangleq R_n^T J_p$$

and

$$\widehat{\underline{R}}_n \triangleq J_{p,n} \underline{R}_n, \qquad \widehat{\underline{R}}_n^T \triangleq \underline{R}_n^T J_{p,n} \quad (27)$$

It follow from (24) $\sim$ (27)

$$\begin{aligned} \underline{W}_{n+1} &= -R_{p,n+1}^{-1} \underline{R}_{n+1} \\ &= -R_{p,n+1}^{-1} J_{p,n+1} \widehat{\underline{R}}_{n+1} \\ &= -J_{p,n+1} R_{p,n+1}^{-T} \widehat{\underline{R}}_{n+1} \end{aligned}$$

$$= -J_{p,n+1} \begin{bmatrix} R_{p,n}^{-T} + \underline{V}_n \alpha_n^T \underline{W}_n^T & \underline{V}_n \alpha_n^{-T} \\ \alpha_n^{-T} \underline{W}_n^T & \alpha_n^{-T} \end{bmatrix} \begin{bmatrix} \widehat{\underline{R}}_n \\ \widehat{R}_{n+1} \end{bmatrix}$$

$$= -J_{p,n+1} \begin{bmatrix} R_{p,n}^{-T} \widehat{\underline{R}}_n + \underline{V}_n \alpha_n^{-T} (\underline{W}_n^T \widehat{\underline{R}}_n + \widehat{R}_{n+1}) \\ \alpha_n^{-T} (\underline{W}_n^T \widehat{\underline{R}}_n + \widehat{R}_{n+1}) \end{bmatrix} \quad (28)$$

Thus, we obtain

$$\widehat{\underline{W}}_{n+1} = \begin{bmatrix} \widehat{\underline{W}}_n \\ \mathbf{0}_p \end{bmatrix} - \begin{bmatrix} \underline{V}_n \\ I_p \end{bmatrix} \alpha_n^{-T} \beta_n \quad (29)$$

where

$$\widehat{\underline{W}}_n = J_{p,n} \underline{W}_n \quad (30)$$

$$\beta_n = \underline{W}_n^T \widehat{\underline{R}}_n + \widehat{R}_{n+1}. \quad (31)$$

Similarly, $\widehat{\underline{V}}_{n+1}$ can be rewritten as

$$\widehat{\underline{V}}_{n+1} = \begin{bmatrix} \widehat{\underline{V}}_n \\ \mathbf{0}_p \end{bmatrix} - \begin{bmatrix} \underline{W}_n \\ I_p \end{bmatrix} \alpha_n^{-1} \gamma_n \quad (32)$$

where

$$\widehat{\underline{V}}_n = J_{p,n} \underline{V}_n \quad (33)$$

$$\gamma_n = \underline{V}_n^T \widehat{\underline{R}}_n^T + \widehat{\underline{R}}_{n+1}^T. \quad (34)$$

The recursion for α_n is obtained from (24) $\sim$ (27) and (33) $\sim$ (35) as follows

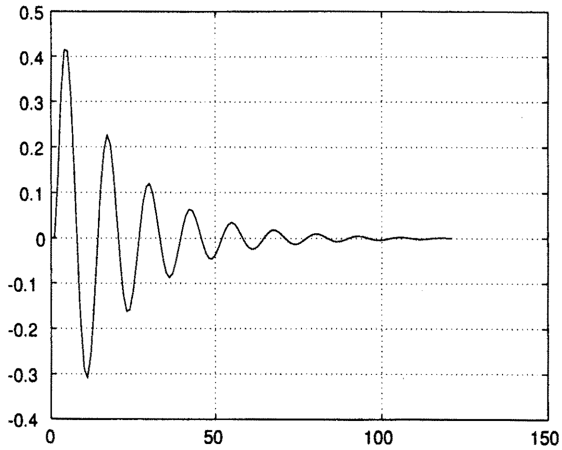

Fig. 1.: Impulse response

$$\alpha_{n+1} = \alpha_n - \gamma_n^T \alpha_n^{-T} \beta_n. \tag{35}$$

Equations (30) $\sim$ (35) and (36) constitute the recursions necessary for carrying out the recursive inversion scheme suggested by (24) $\sim$ (27).

5. Example

We give a simple example to illustrate the advantage of using Kautz function for second order system. Consider a contiuous time transfer function

$$G^0(s) = \frac{1}{s^2 + 0.2s + 1} \tag{36}$$

with resonant frequency $\omega_0 = 1$ and damping 0.1. This system is sampled using a zero-order hold with sampling period $T = 0.5$.
Impulse response of system is **Fig. 1**. **Fig. 2** is the result of FIR model where the order is 100. The choice $a = 0.84$, $n = 12$, which corresponds to the real part of the poles of the true system, is illustrated in **Fig. 3**, **4**. Let use take $b = 0.91$ and $c = -0.92$. The Kautz approximation of order 6 is shown in **Fig. 5**, **6**. And a **Fig. 5** is the result that it was calculated by using the Block-Toeplitz matrix.

6. Conclusion

In this paper we have analysed some asymptotic properties of linear estimation schemes that identify a finite number of expansion coefficients in a series expansion of a linear stable transfer function, employing recently developed generalized orthogonal basis functions. The basis functions generalize the

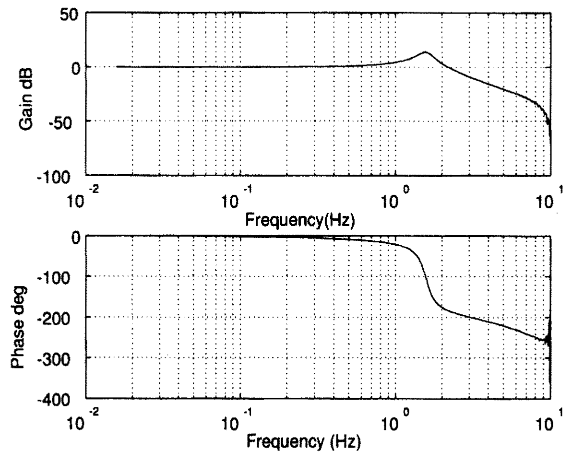

Fig. 2. Bode plots: Solid line-true system, dashed line-FIR model of order 100

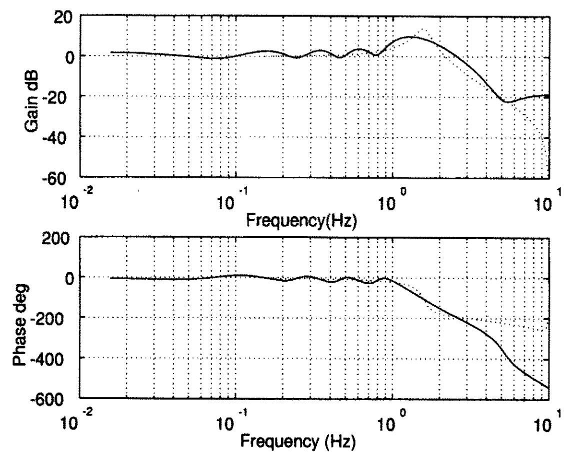

Fig. 3. Bode plots: Solid line-true system, dashed line-Laguerre model of order 12

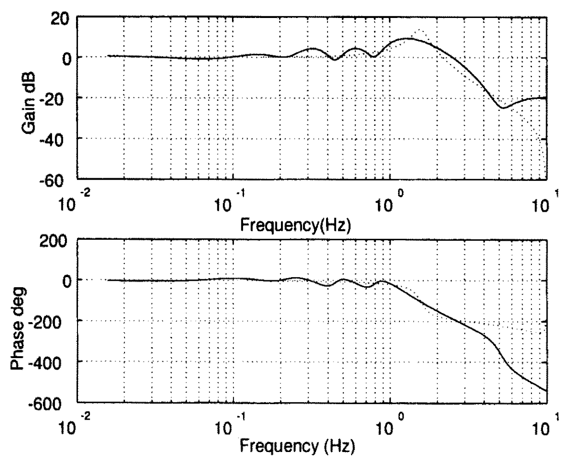

Fig. 4. Bode plots: Solid line-true system, dashed line-Laguerre model of order 12

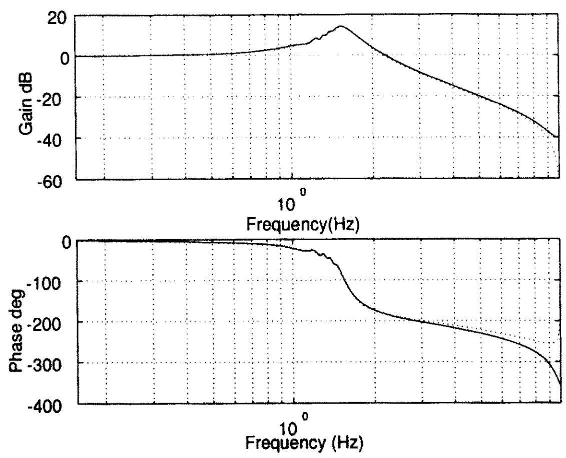

Fig. 5. Bode plots: Solid line-true system, dashed line-Kautz model of order 6

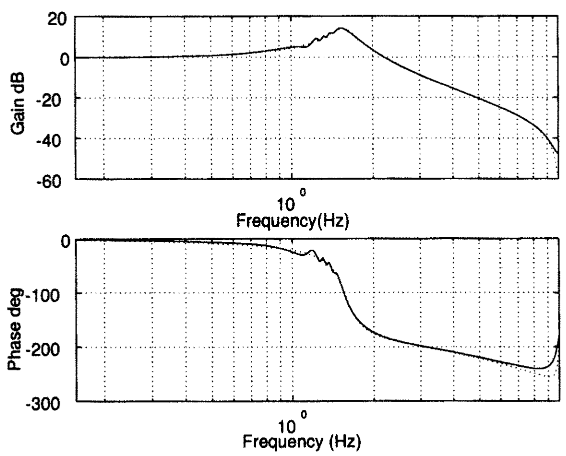

Fig. 6. Bode plots: Solid line-true system, dashed line-Kautz model of order 6

well known pulse, Laguerre and Kautz basis functions. We illustrated by numerical example that the presented method of identification can be performed with good accuracy using a rather smaller numbers of expansion terms than that for the case where the FIR model is used.

References

1) L. Ljung, "System Identification, Theory for the User," *Prentice Hall*, NJ, 1987.
2) C. Hwang and Y. Shih, "Parameter identification via Laguerre polynomials, " *Int. J. Systems Sci.*, vol. 13, no. 2, pp. 209–217, 1982.
3) B. Wahlberg, "System identification using Laguerre models, " *IEEE Trans. Autom. Control*, vol. 36 no. 5, pp. 551–562, 1991.
4) B. Wahlberg, "System identification using Kautz models, " *IEEE Trans. Autom. Control*, vol. 39, no. 6, pp.1276–1282, 1994.
5) Y. Fu and G. A. Dumont, "An optimum time scale for discrete Laguerre network, " *IEEE Trans. Autom. Control*, vol. 38, no. 6, pp. 934–938, 1993.
6) P. R. Clement, "Laguerre functions in signal analysis and parameter identification, " *J. Franklin Inst*, vol. 313, no. 2, pp. 85–95, 1982.
7) Y. W. Lee, "Choice of the time-scaling factor for linear system approximations using orthonormal Laguerre functions, " *IEEE Trans. Autom. Control*, pp. 487–489, 1965.
8) Y. W. Lee "Statistical theory of communication", *New York: Wiley*, 1960.
9) M. Wax and T. Kailath, "Efficient inversion of Toeplitz-Block Toeplitz matrix," *IEEE Trans. Acous. Speech*, pp. 1218–1221, 1983.
10) N. Kalouptsidis and S. Theodoridis, "Adaptive system identification and signal processing algorithms," *Prentice Hall*, 1993.
11) P. M. J. Van den Hof et. al. "System identification with generalized orthonormal basis functions", *Automatica*, vol. 31, no. 12, pp. 1821–1834, 1995.