Available online at www.sciencedirect.com
SCIENCE DIRECT®

Automatica 41 (2005) 1905–1916

automatica

www.elsevier.com/locate/automatica

Strong consistence of recursive identification for Wiener systems☆

Xiao-Li Hu[a, b], Han-Fu Chen[a, *]

[a]Key Laboratory of Systems and Control, Institute of Systems Science, AMSS Chinese Academy of Sciences, Beijing 100080, PR China
[b]Graduate School of the Chinese Academy of Sciences, Beijing 100039, PR China

Received 17 December 2004; received in revised form 18 April 2005; accepted 11 June 2005

Abstract

The paper concerns identification of the Wiener system consisting of a linear subsystem followed by a static nonlinearity $f(\cdot)$ with no invertibility and structure assumption. Recursive estimates are given for coefficients of the linear subsystem and for the value $f(v)$ at any fixed v. The main contribution of the paper consists in establishing convergence with probability one of the proposed algorithms to the true values. This probably is the first strong consistency result for this kind of Wiener systems. A numerical example is given, which justifies the theoretical analysis.

Keywords: Wiener system; Nonparametric nonlinearity; Recursive estimate; Strong consistency; Stochastic approximation

1. Introduction

The Hammerstein and Wiener systems, in particular, their identification issue have attracted a great attention from researchers because of their importance in applications. Since these systems are nonlinear, the identification methods demonstrated by Chen & Guo (1991) and Ljung (1987) are not directly applicable.

A linear system cascaded with a static nonlinearity is called the Wiener (or Hammerstein) system if the nonlinearity follows (or is followed by) the linear subsystem. This paper concerns with identification of the SISO Wiener system presented in Fig. 1 where u_k is the one-dimensional system input to be designed, v_k is the output of the linear subsystem serving as the input of the memoryless nonlinear block, and y_k is the system output which is observed with additive noise ε_k. The coefficients of the linear subsystem and the nonlinear function $f(\cdot)$ are unknown. The problem is how to estimate coefficients contained in the linear subsystem and the static nonlinearity $f(\cdot)$ on the basis of observation $\{z_k\}$ and the adequately designed input $\{u_k\}$, where

$$z_k = y_k + \varepsilon_k. \tag{1}$$

The name Wiener model probably comes from the famous book by Wiener (1958), where the nonlinearity is expanded to the functional series and the correlation analysis is carried out by using the Gaussian input. Based on the method proposed by Wiener (1958) there were many works on analysis and identification of nonlinear systems in 1960s and 1970s. Among early works on identification of Wiener systems, a practical nonparametric algorithm is proposed by Billings & Fakhouri (1978) where no inversion of the nonlinearity is required.

For characterizing the nonlinearity the parametric approach (Bendat, 1999; Hasiewicz, 1987; Hunter & Korenberg, 1986; Nordsjö & Zetterberg, 2001; Pajunen, 1992; Verhaegen & Westwick, 1996; Vörös, 2001; Westwick & Kearney, 1992; Wigren, 1993, 1994) is mostly applied in literature, but the nonparametric approach is also considered (Billings & Fakhouri, 1978; Greblicki, 1997, 2001).

☆This paper was not presented at any IFAC meeting. This paper was recommended for publication in revised form by Associate Editor Johan Schoukens under the direction of Editor Torsten Soederstroem. This work is supported by NSFC (Projects G0221301, 60334040, 60474004.
* Corresponding author. Tel.: +86 10 62579540; fax: +86 10 62587343.
E-mail addresses: xlhu@amss.ac.cn (X.-L. Hu), hfchen@iss.ac.cn (H.-F. Chen).

0005-1098/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.automatica.2005.06.006

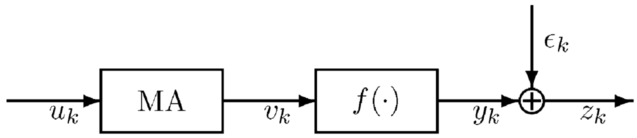

Fig. 1. Wiener system.

When the parametric approach is applied, the nonlinearity is presented either as a linear combination of known functions with unknown coefficients (Hasiewicz, 1987; Hunter & Korenberg, 1986; Nordsjö & Zetterberg, 2001; Westwick & Kearney, 1992) or as a piecewise linear function (Pajunen, 1992; Vörös, 2001; Wigren, 1993). In this case, the parameter estimates may be derived by minimizing some specially designed loss function, and this can be realized by using any optimizing algorithm for data with fixed sample size. Proceeding in this way, the parameters cannot be updated online as can be seen in Vörös (2001). Nevertheless, the estimates may still be made recursive and even with certain kind of convergency, if rather restrictive conditions are imposed as demonstrated by Wigren (1993, 1997, 1998) the nonlinear function is assumed to be known.

When the nonparametric approach is considered, the nonlinear function is usually required to be invertible (Greblicki, 1997, 2001), and the argument v for any given $u = f(v)$ rather than $f(v)$ for any given v is estimated. This may limit applications of corresponding identification methods in practice by the following consideration: saturations are not invertible, but they quite often exist in practical systems and affect the measured outputs; also, inversion of the nonlinearity can lead to severe amplification of possible measurement disturbances as pointed out by Wigren (1993), etc.

The goal of this paper is to recursively estimate the coefficients of the linear subsystem and the value $f(v)$ for any given v without requiring invertibility of $f(\cdot)$. The estimates are required to be strongly consistent, i.e., to converge to the true values with probability one. A similar problem for Hammerstein systems is solved by Chen (2004) by using stochastic approximation (SA) algorithms with expanding truncations (Chen, 2002). There the input is designed to be a sequence of bounded iid random variables, and $f(\cdot)$ is estimated with the help of a kernel function applied to the SA algorithm.

Let us explain why SA is an appropriate tool to deal with the identification problem. When estimating an unknown parameter ϑ on the basis of observation data denoted by $\{g_k\}$, one can always transform this to a SA problem, i.e., to a root-seeking problem for any function $g(\cdot)$ with root ϑ, e.g., $g(x) = -(x - \vartheta)$. This is because g_{k+1} can always be written as $g_{k+1} = g(x_k) + \eta_{k+1}$ with $\eta_{k+1} \triangleq g_{k+1} - g(x_k)$, where x_k denotes the kth estimate for ϑ. In other words, the observation data $\{g_k\}$ can be viewed as a noisy observation on $g(x_k)$ with additive noise η_k.

It is natural to come to the idea: to solve the stated problem for Wiener systems by using SA algorithms with expanding truncations and with kernel functions. However, in doing so, there is an essential difference in analysis for Wiener systems from that for Hammerstein systems. To explain this, we note that the analysis given by Chen (2004) is essentially based on two facts: (1) The correlation function between the input and output of the system has a simple analytic expression connecting parameters to be estimated; (2) All signals passing through the system are bounded when the input is bounded. As shown by Chen (2004), for Hammerstein systems a sequence of bounded iid random variables serving as the system input results in these two properties.

For Wiener systems, though a bounded input still implies the boundedness of all signals in the system, the correlation function between the input and output of the system, in general, does not have a simple analytic expression. This hints us to take an iid Gaussian random variables to serve as the system input. However, the Gaussian random variable is unbounded, and hence the Gaussian input may give rise to the unboundedness of signals in the system. This explains why the analysis method given by Chen (2004) cannot directly be applied to the present case.

The requirement for boundedness of signals passing through the system can also be explained by the following intuitive observation. To estimate $f(v)$ it is important to recover the input v_k of the nonlinear function. The estimate for v_k, denoted by $\hat{v}_k$, is obtained as the output of the estimated linear subsystem, which means the subsystem with coefficients replaced by their estimates. However, $\hat{v}_k$ may not be close to v_k even if the estimates for coefficients of the linear subsystem are sufficiently accurate, when $\{u_k\}$ is unbounded.

To overcome this difficulty, we proceed as follows. While the system input $\{u_k\}$ is taken to be a sequence of iid Gaussian random variables, not all u_k but only such $r + 1$ successive u_k that are bounded by a given constant are used to estimate v_k, where r is the order of the linear subsystem. This selection guarantees that $\{v_k\}$ generated by sets of $r + 1$ successive bounded u_k is bounded. Since the selection depends on sample paths, we have to use the concept of stopping time, which is well developed in probability theory (see, e.g., Chow & Teicher (1978)).

The rest of the paper is organized as follows. The system considered in the paper and conditions imposed on the system are given in Section 2. Also, the basic results of SA used in the paper are described in Section 2. The recursive identification algorithms and their strong consistency for estimating the linear subsystem and the nonlinear block are, respectively, presented in Sections 3 and 4. A numerical example is demonstrated in Section 5 and some concluding remarks are given in Section 6. The mathematical details concerning the properties of stopping times and the behaviors of kernel functions are given in Appendix.