NONPARAMETRIC APPROACH TO IDENTIFICATION OF BLOCK-ORIENTED SYSTEMS

Grzegorz Mzyk, Zygmunt Hasiewicz

The Institute of Computer Engineering, Control and Robotics
Wrocław University of Technology
Janiszewskiego 11/17, 50-372 Wrocław, Poland
tel. +48 71 320 25 49, fax. +48 71 321 26 77
e-mail: Grzegorz.Mzyk@pwr.wroc.pl

Abstract: An alternative philosophy of stochastic nonlinear system modelling is presented. We analyse the problem of the static nonlinearity, and the linear dynamics estimation in complex systems, under conditions of poor a priori knowledge. The parametric description of the other component and correlation structure of signals are assumed to be completely unknown. Two kinds of nonparametric methods (kernel regression and orthogonal expansion) are reconsidered from the point of view of their application for supporting standard methods (least squares, instrumental variables). The most important asymptotic properties of the proposed estimates are given.

Keywords: System identification, nonparametric methods, kernel regression, orthogonal expansion, wavelets, least squares, instrumental variables.

1. INTRODUCTION

Identification routine is in general based on two kinds of information. Firstly, we have at own disposal a priori expert knowledge in the closed formula, describing the system structure, with finite number of unknown parameters. Secondly, we have the set of input-output measurements (learning sequence). Unfortunately, if the model assumed at the beginning is not correct the systematic approximation error appears. Here we consider the nonparametric identification methods (Greblicki, and Pawlak, 1986; Greblicki, and Pawlak, 1994; Greblicki, 2004; Pawlak, and Hasiewicz, 1998) which do not involve any a priori knowledge and are based on measurements only. We analyse possibilities of their application for supporting the traditional least squares and instrumental variables methods in tasks with poor a priori knowledge of the random disturbances and under given parametric formulas describing both linear and nonlinear components of the complex system (Hasiewicz, and Mzyk, 2004; Hasiewicz, and Mzyk, 2006). Results of the illustrative computer simulation example are also included.

2. BLOCK-ORIENTED SYSTEMS

The conception of block-oriented models had been commonly accepted in eighties. It means that the nonlinear transformation may be represented as interconnections of nonlinear static elements and linear dynamics. The most popular structures are the Hammerstein (Fig. 1) system and the Wiener system. They have good capabilities of modelling of many real processes (i.e. nonlinearity recovering in electric motors, noise-cancellation systems, modelling of heat exchange processes).

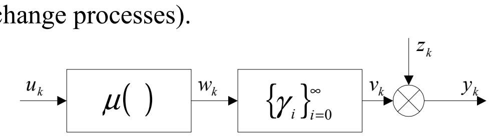

Fig. 1. Hammerstein system.

We assume that the input u_k is an i.i.d. random sequence, the linear subsystem is asymptotically stable, the output noise z_k is zero-mean, ergodic and independent of input. The characteristic $\mu()$ is *any* nonlinear function. Only the input u_k and the output y_k ($k = 1,2,...,M$) of the whole system can be measured, the signals w_k and v_k are not accessible. The aim is to estimate $\mu(u)$ for a given point u.

3. NONPARAMETRIC METHODS

Since the regression function in Hammerstein system

$$R(u) = \mathbf{E}\{y_k \mid u_k = u\} = c\mu(u) + s \quad (1)$$

is scaled and shifted version of the system nonlinearity, $\mu()$ may be estimated only up to two unknown constants c and s, independently of the identification method. It is consequence of the complex structure of the system (inaccessibility of w_k). Therefore, without any loss of generality assume, for clarity of exposition, that $\gamma_0 = 1$ and $\mu(0) = 0$. Then we obtain $R(u) = \mu(u)$.

3.1 Kernel regression estimation

Nonparametric kernel estimate has the form

$$\hat{\mu}_M(u) = \sum_{i=1}^{M} y_i K\left(\frac{u-u_i}{h(M)}\right) \Big/ \sum_{i=1}^{M} K\left(\frac{u-u_i}{h(M)}\right) \quad (2)$$

where $K()$ is a kernel function, and $h(M)$ − band-width parameter. The following theorems hold.

Theorem 1. If $h(M) \to 0$ and $Mh(M) \to \infty$ as $M \to \infty$ then

$$\hat{\mu}_M(u) \to \mu(u) \quad (3)$$

in probability as $M \to \infty$ in each continuity point of $\mu()$ and the input probability density.

Theorem 2. If $\mu()$ is twice differentiable in the point u then for $h(M) = O(M^{-1/5})$ it holds that

$$\left|\hat{\mu}_M(u) - \mu(u)\right| = O(M^{-2/5}) \quad (4)$$

3.2 Orthogonal expansion

Observe that (Greblicki, and Śliwiński, 2001)

$$\mu(u) = \frac{g(u)}{f(u)}$$

where $g(u) = \mu(u) f(u)$ and $f(u)$ is the input probability density function. Let $\phi_i(u)$, $i = 0,1,2,...$ be a sequence of complete orthonormal functions. We can expand $g()$ and $f()$ as follows

$$g(u) = \sum_{i=0}^{\infty} a_i \phi_i(u), \quad f(u) = \sum_{i=0}^{\infty} b_i \phi_i(u)$$

where

$$a_i = \mathbf{E} y_k \phi_i(u_k), \quad b_i = \mathbf{E}\phi_i(u_k)$$

may be simply estimated in the following way

$$\hat{a}_i = \frac{1}{M}\sum_{k=1}^{M} y_k \phi_i(u_k), \quad \hat{b}_i = \frac{1}{M}\sum_{k=1}^{M} \phi_i(u_k)$$

The nonlinearity estimate has the form

$$\hat{\mu}_M(u) = \frac{\sum_{i=0}^{q(M)} \hat{a}_i \phi_i(u)}{\sum_{i=0}^{q(M)} \hat{b}_i \phi_i(u)}$$

where $q(M) \to \infty$ as $M \to \infty$, but for assuring the convergence of $\hat{\mu}_M(u)$ to $\mu(u)$, the rate of $q(M)$-increasing must be appropriately slow (see Table 1).

Table 1 The convergence conditions

$\{\phi_i(u)\}$	the condition
trigonometric series	$\lim_{M\to\infty} \frac{q^2(M)}{M} = 0$
Legendre series	$\lim_{M\to\infty} \frac{q^2(M)}{M} = 0$
Laguerre series	$\lim_{M\to\infty} \frac{q^6(M)}{M} = 0$
Hermite series	$\lim_{M\to\infty} \frac{q^{5/3}(M)}{M} = 0$
Daubechies wavelets	$\lim_{M\to\infty} \frac{2^{2q(M)+2}}{M} = 0$

4. COMBINED PARAMETRIC-NONPARAMETRIC ALGORITHMS

A novel is application of *plug-in* technique and inserting the nonparametric estimates $\hat{w}_{k,M} = \hat{\mu}_M(u)$ instead of the interaction signal w_k into the traditional standard least squares or instrumental variables formula involving N first measurements (see Hasiewicz, and Mzyk, 2004).

4.1 Estimation of the static nonlinearity

The generalised least squares estimate has the following form

$$\hat{c}_{N,M} = \left(\Phi_N^T \Phi_N\right)^{-1} \Phi_N^T \hat{W}_{N,M} \quad (5)$$

Here we assume that

$$\mu(u_k) = \phi^T(u_k)c \tag{6}$$

i.e. the nonlinear characteristic is linear in parameters,

$$\Phi_N = (\phi_1(u_k), \phi_2(u_k), ..., \phi_N(u_k))^T \tag{7}$$

where

$$\phi(u_k) = \left(f_1(u_k), f_2(u_k), ..., f_m(u_k)\right)^T \tag{8}$$

$f_1(), f_2(), ..., f_m()$ is a set of a priori known basis functions,

$$\hat{W}_{N,M} = (\hat{w}_{1,M}, \hat{w}_{2,M}, ..., \hat{w}_{N,M})^T \tag{9}$$

and

$$c = (c_1, c_2, ..., c_m)^T \tag{10}$$

is vector of unknown parameters. Similar approach is applicable also for the nonlinearities of the general form $\mu(u_k) = \mu(u_k, c)$ (e.g. $\mu(u_k) = e^{c_1 u} + \sin(c_2 u_k)$) by minimization of the empirical criterion

$$\hat{c}_{N,M} = \arg\min_c \sum_{k=1}^{N} \left(\hat{w}_{k,M} - \mu(u_k, c)\right)^2 \tag{11}$$

Thanks to nonparametric estimation of the interaction signal w_k we obtain full decomposition of the identification task into two problems. It allows us to apply Levenberg-Marquardt method, simulated annealing technique, or genetic algorithms to solve (11).

4.2 Identification of the linear dynamics

Analogously, for the linear subsystem we propose the combined parametric-nonparametric routine

$$\hat{\theta}_{N,M} = \left(\hat{\Psi}_{N,M}^T \hat{\Theta}_{N,M}\right)^{-1} \hat{\Psi}_{N,M}^T Y_N \tag{12}$$

where

$$Y_N = (y_1, y_2, ..., y_N)^T, \; y_k = \vartheta_k^T \theta + z_k . \tag{13}$$

The vector

$$\theta = (\alpha_0, \alpha_1, ..., \alpha_s, \beta_1, \beta_2, ..., \beta_p)^T \tag{14}$$

consists of unknown parameters of linear dynamics,

$$\hat{\Theta}_{N,M} = (\vartheta_1, \vartheta_2, ..., \vartheta_N)^T, \tag{15}$$

where

$$\vartheta_k = (w_k, w_{k-1}, ..., w_{k-s}, y_{k-1}, y_{k-2}, ..., y_{k-p})^T \tag{16}$$

is generalised input matrix and $\hat{\Psi}_{N,M}$ is instrumental variables matrix, generated in nonparametric way.

All algorithms presented in Section 4 have strict convergence proofs (Hasiewicz, and Mzyk, 2004; Hasiewicz, and Mzyk, 2006). It was also proved that the optimal instruments have the form

$$\psi_k^* = (w_k, ..., w_{k-s}, \bar{y}_{k-1}, ..., \bar{y}_{k-p})^T \tag{17}$$

where $\bar{y}_k$'s are noise-free outputs, which are inaccessible for a direct measurements. We can approximate it

$$\psi_k^\# = (w_k, ..., w_{k-s}, \tilde{y}_{k-1}, ..., \tilde{y}_{k-p})^T \tag{18}$$

with use of the standard correlation method

$$\tilde{y}_k = \sum_{i=0}^{APR} \hat{\gamma}_{i,M} \hat{w}_{k-i,M} \tag{19}$$

where

$$\hat{\gamma}_{i,M} = \hat{\chi}_{i,M} / \hat{\chi}_{0,M} \tag{20}$$

$$\hat{\chi}_{i,M} = \frac{1}{M} \sum_{k=1}^{M-i} (y_{k+i} - \bar{y})(u_k - \bar{u}) \tag{21}$$

$$\bar{y} = \frac{1}{M} \sum_{k=1}^{M} y_k, \quad \bar{u} = \frac{1}{M} \sum_{k=1}^{M} u_k \tag{22}$$

4. SIMULATION EXAMPLE

As the illustration, we simulated a simple Hammerstein system excited by the i.i.d. random sequence of uniform distribution

$$u_k \sim U[-1,1] \tag{23}$$

The nonlinear characteristic was the polynomial of the known order $m = 3$

$$\mu(u_k) = c_1 u_k + c_2 u_k^2 + c_3 u_k^3 \tag{24}$$

and the true but unknown parameters were set as follows

$$c_1 = 1, \quad c_2 = 1, \quad c_3 = -1 \tag{25}$$

Under requirement of stability of the *IIR* linear filter we put

$$\gamma_i = 2^{-i}, \quad v_k = \sum_{i=0}^{\infty} 2^{-i} w_{k-i} \tag{26}$$

Equation (20) can be written in the equivalent, parametric form

$$v_k = a_0 w_k + b_1 v_{k-1} \tag{27}$$

where $a_0 = 1$, and $b_1 = 1/2$ are unknown parameters we need to estimate.
The coloured additive output noise were generated according to the rule

$$z_k = \varepsilon_k + \varepsilon_{k-1} + 0.7 z_{k-1} \tag{28}$$

where $\varepsilon_k \sim U[-0.1, 0.1]$. Estimation has been performed for the learning sequence of the length $M = 500$ ($N = 100$). The kernel regression estimate of the form given by (2) has been applied with the Gaussian kernel function

$$K(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2} \tag{29}$$

and with the bandwidth parameter

$$h(M) = M^{-1/5} \tag{30}$$

which fulfils the assumptions of Theorem 1 and Theorem 2. Then the parametric knowledge given by (24) has been used for evaluation of the estimate (5). Results are shown in Fig. 2.

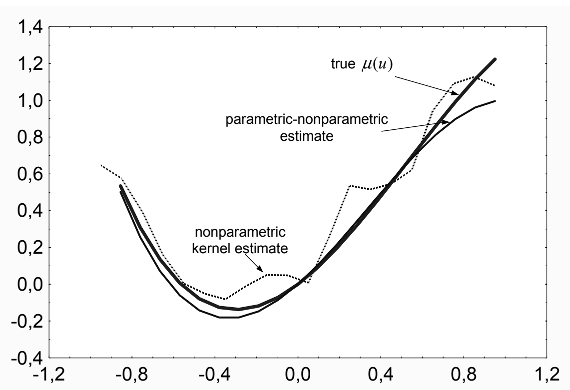

Fig. 2. Results of nonlinearity estimation.

For modelling of the linear subsystems the cross-correlation analysis has been performed first (see (20)-(22) and the results on Fig. 3). Next, the parametric knowledge given by (27) has been involved in (12) with $APR = 3$. In consequence, the parameter estimates $\hat{a}_0 = 0.998$, and $b_1 = 0.514$ have been provided.

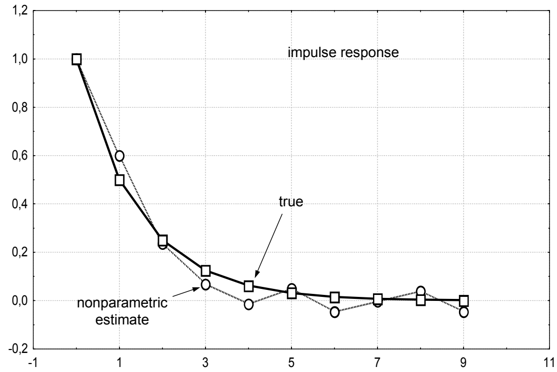

Fig. 3. The impulse response and its estimate.

5. CONCLUSIONS

The problems of identification of complex systems with nonlinearities, which are not linear in parameters disturbed by coloured noise belongs to the most difficult class of tasks and in opinion of authors has not been so far considered in literature.

REFERENCES

Greblicki, W. and M. Pawlak (1986). Identification of discrete Hammerstein systems using kernel regression estimates. *IEEE Transactions on Automatic Control*, vol. 31, pp. 74-77.

Greblicki, W. and M. Pawlak (1994). Cascade non-linear system identification by a non-parametric method. *International Journal of System Science*, vol. 25, pp. 129-153.

Greblicki, W. and P. Śliwiński (2001). Orthogonal series algorithms to identify Hammerstein systems, *Proceedings of the 7th IEEE International Conference on Methods and Models in Automation and Robotics*, Międzyzdroje, Poland, pp. 1009-1014.

Greblicki, W. (2004). Nonlinearity recovering in Wiener system driven with correlated signal, *IEEE Transactions on Automatic Control*, vol.49, pp. 1805-1809.

Hasiewicz, Z. and G. Mzyk (2004). Combined parametric-nonparametric identification of Hammerstein systems. *IEEE Transactions on Automatic Control*, vol. 49, pp. 1370-1376.

Hasiewicz, Z. and G. Mzyk (2006). Nonparametric instrumental variables for Hammerstein system identification. *submitted for publication in the IEEE Transactions on Automatic Control.*

Pawlak, M. and Z. Hasiewicz (1998). Nonlinear system identification by the Haar multiresolution analysis. *IEEE Transactions on Circuits and Systems*, vol. 45, pp. 945-961.