Applied Mathematical Modelling 38 (2014) 2414–2421

Contents lists available at ScienceDirect

Applied Mathematical Modelling

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

Volterra-type models for nonlinear systems identification

C.A. Schmidt, S.I. Biagiola *, J.E. Cousseau, J.L. Figueroa

Instituto de Investigaciones en Ingeniería Eléctrica (UNS – CONICET), Avda. Alem 1253, 8000 Bahía Blanca, Argentina

ARTICLE INFO

Article history:
Received 16 July 2012
Received in revised form 12 September 2013
Accepted 8 October 2013
Available online 1 November 2013

Keywords:
Nonlinear identification
Volterra-type models
Wiener model
Hammerstein model

ABSTRACT

In this work, multi-input multi-output (MIMO) nonlinear process identification is dealt with. In particular, two Volterra-type models are discussed in the context of system identification. These models are: Memory Polynomial (MP) and Modified Generalized Memory Polynomial (MGMP), which can be considered as a generalization of Hammerstein and Wiener models, respectively. Both of them are appealing representations as they allow to describe larger model sets with less parametric complexity. Simulation example is given to illustrate the quality of the obtained models.

1. Introduction

Several model-based nonlinear controller design techniques have been developed in recent decades. Since most of them require models of relatively low complexity, use of first principle models is not an alternative to control design. In this context, empirical modeling is a useful option, because it allows a much greater influence on the model complexity. This influence comes from the fact that before a nonlinear model of a plant is identified, a model structure has to be chosen. Moreover, this model should be capable of representing the dynamical behavior of the plant [1].

The problem of finding a good and simple nonlinear model structure is a complex issue. It is well-known that systems presenting weak nonlinearities could be represented by a Volterra model. In particular, discrete-time systems with fading memory can be approximated arbitrarily well by a discrete-time Volterra model (DTVM), if adequate orders are chosen [2–4].

In general, the most general form of nonlinearity with M-tap memory (i.e., M tap delay lines or memory effects) is described as follows [5]:

$$\hat{y}(t) = \sum_{n=1}^{N} v_M^n(t), \tag{1}$$

with $\hat{y} \in \Re$ and where

$$v_M^n(t) = \sum_{i_1=0}^{M-1} \cdots \sum_{i_n=0}^{M-1} h_n(i_1, \ldots, i_n) \prod_{l=1}^{n} u(t - i_l) \tag{2}$$

* Corresponding author. Tel.: +54 291 4595100x3338.
E-mail addresses: cschmidt@uns.edu.ar (C.A. Schmidt), biagiola@uns.edu.ar (S.I. Biagiola), jcousseau@uns.edu.ar (J.E. Cousseau), figueroa@uns.edu.ar (J.L. Figueroa).

0307-904X/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.apm.2013.10.041

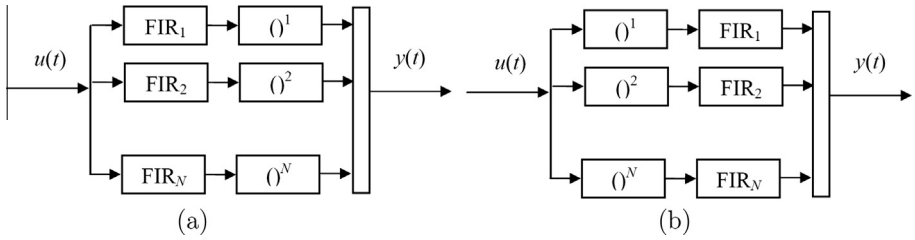

Fig. 1. Modified Generalized Memory Polynomial models (a) and Memory Polynomial (b).

is the n-dimensional convolution of the input $u \in \Re$ with the Volterra kernel h_n. This model, which is linear in the parameters, allows the convergence to a global optimum for any identification algorithm. However, the main drawback of this approach is the large number of parameters involved, giving rise to a high complexity model.

To overcome this problem, it is a common practice to appeal to some simplified DTVM models. For example, the use of Wiener or Hammerstein models are very frequent in many application fields ([6] and references therein). Although both Wiener and Hammerstein models are quite simple structures, they show a limitation as for the class of systems they can represent.

Recently, two appealing versions of the DTVM model have emerged in the context of compensation of nonlinear systems. Simplicity and generalization are their two outstanding properties. They are the Memory Polynomial (MP) [7,8] and the Modified Generalized Memory Polynomial (MGMP) [9]. These models allow the representation of a much more complex type of systems than Hammerstein and Wiener ones. In their typical form, these models adopt the structure depicted in Fig. 1.

The MP model is a generalization of a Hammerstein system [7], where multiple linear filters can be included after the static nonlinearity. Provided that the structure is formed by an Nth order polynomial and a finite impulse response (FIR) filter of order M, then the model output is given by

$$\hat{y}^{MP}(t) = \sum_{n=1}^{N}\left(\sum_{m=0}^{M-1}\alpha_{nm}u^{n}(t-m)\right), \tag{3}$$

where α_{nm} (with $n = 1,\ldots,N$ and $m = 1,\ldots,M$) stands for the compensator parameters. This structure has the advantage that the output signal is linear in the unknown parameters (i.e., the terms α_{nm}), which renders efficient parameter estimation which can be straightforwardly performed through least-squares methods.

An alternative is the MGMP model, it can be shown that in this case, the input–output relationship is given by

$$\hat{y}^{MGMP}(t) = \sum_{n=1}^{N}\left(\sum_{m=0}^{M-1}\alpha_{nm}u(t-m)\right)^{n}. \tag{4}$$

This model is the transpose of the block diagram presented for the MP model, with the power terms after the FIR filter in each parallel branch. Note that this model includes cross-terms among the samples $u(t-m)$. This is due to the fact that for a given n, the power terms in brackets at the left-hand side of Eq. (4) can be phrased as

$$\left(\sum_{m=0}^{M-1}\alpha_{nm}u(t-m)\right)^{n} \propto \sum_{m_1=0}^{M-1}\sum_{m_2=m_1}^{M-1}\cdots\sum_{m_n=m_{n-1}}^{M-1}\alpha_{nm_1}\cdots\alpha_{nm_n}u(t-m_1)\cdots u(t-m_n). \tag{5}$$

Both MP and MGMP models are special cases of finite Volterra models [3,7]. The model described by Eq. (5) is a generalization of a Wiener model different to the GMP derived in [7]. Note that the MGMP model output is not linear in the model parameters, this will lead, in general, to a complex identification algorithm.

A disadvantage of these models, based on FIR descriptions to represent dynamics behavior, is that they can demand too many parameters to describe those systems whose impulse response decays slowly. To cope with this fact, we propose the use of linear filters which incorporate prior knowledge about the process dynamics in order to reduce the number of parameters. Specifically, the FIR description is replaced by orthonormal Laguerre or Kautz filters.

The rest of the paper is organized as follows. In Section 2, the proposed model is defined, and the identification algorithms are presented. Section 3 states a discussion on the application of the proposed modeling approach for the identification of a distillation column. Finally, the paper ends with some conclusions in Section 4.

2. Volterra-type models

In this section, an alternative model is presented to represent multivariable nonlinear systems. The main underlying idea is to replace the tap delay used by the FIR filters depicted in Fig. 1 by an orthogonal basis [10,11].

Fig. 2 depicts the internal structure of the ith FIR filter shown in Fig. 1. On the other hand, Fig. 3 shows the internal structure when the proposed Laguerre bases replace the FIR filters.

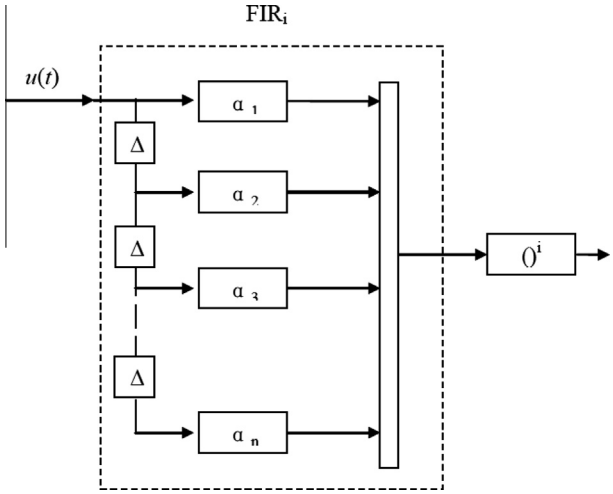

Fig. 2. Configuration of the FIR structure.

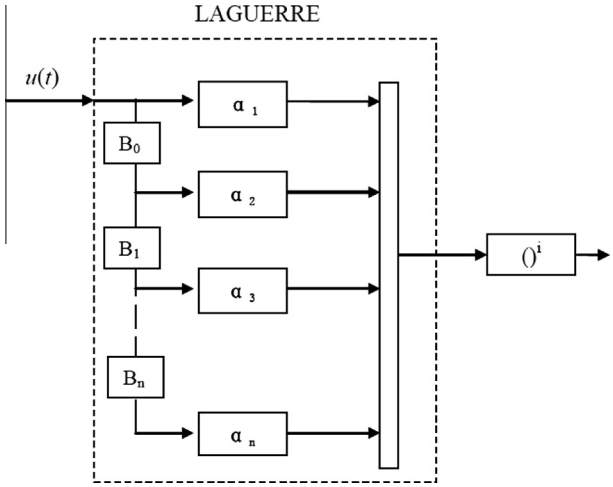

Fig. 3. Configuration of the Laguerre structure.

In this way, the input sequence $\{u(t)\} \in \Re^{N_u}$ is mapped to a sequence of intermediate signals $\{z(t)\} \in \Re^{N_z}$:

$$z_{i,j,p_{\{i,j\}}}(t) = B_{i,j,p_{\{i,j\}}}(q)u_j(t) \tag{6}$$

for $i = 1,\ldots,N_y$, $j = 1,\ldots,N_u$ y $p_{\{i,j\}} = 1,\ldots,N_{z\{i,j\}}$; where q is the forward time operator, $N_{z\{i,j\}}$ is the number of terms in the orthonormal basis from the jth input to the ith output, u_j is the jth entry of the input vector and the $B_{i,j,p_{\{i,j\}}}(q)$ are the elements that relate the jth input to the ith output via the intermediate variables $z_{i,j,p}$. Therefore, the total number of subsystems is $N_u \cdot N_y$.

The bases are defined as

$$B_{i,j,0}(q) = \frac{(1-\xi_{i,j}^2)^{1/2}}{q-\xi_{i,j}} \tag{7}$$

and

$$B_{i,j,p_{\{i,j\}}}(q) = B_{i,j,(p_{\{i,j\}}-1)}(q)\left(\frac{1-\xi_{i,j}q}{q-\xi_{i,j}}\right), \quad p_{\{i,j\}} = 1,\ldots,N_{z\{i,j\}}. \tag{8}$$

Note that from this definition, the total number of internal variables is

$$N_z = \sum_{i=1}^{N_y}\sum_{j=1}^{N_u} N_{z\{i,j\}}. \tag{9}$$

This model allows to use the previous knowledge about the dominant modes of any of the subsystems from each input to each output. Such knowledge about such modes can be incorporated to the model as parameters $\xi_{i,j}$.

Let us define the components of the vector $z(t)$ that affect the ith output, as follows:

$$z_i(t) = C_i z(t), \tag{10}$$

where $i = 1, \ldots, N_y$ and C_i is a matrix formed by zeros and a single 1 in each file according to the positions of $z_{i,j,p_{\{i,j\}}}(t)$ for $j = 1, \ldots, N_u$ and $p_{\{i,j\}} = 1, \ldots, N_{z\{i,j\}}$. Then, $C_i \in \Re^{N_{z_i} \times N_z}$ with

$$N_{z_i} = \sum_{j=1}^{N_u} N_{z\{i,j\}}. \tag{11}$$

Note that in the representation of this basis, for each input vector $u_j(t), z_{i,j}(t)$ is a real vector.

Given this description in terms of the basis elements, it is possible to define two types of models from the point of view of the form in which $z_i(t) \in \Re^{N_{z_i}}$ is mapped to the output $y_i(t) \in \Re$.

2.1. MP model

In this case, and in concordance with the Fig. 1(b), the model output is

$$\hat{y}_i^{MP}(t) = \sum_{n=1}^{N} \boldsymbol{\alpha}_{ni}^T z_i^n(t), \tag{12}$$

where $\boldsymbol{\alpha}_{ni} \in \Re^{N_{z_i}}$ is the parameter vector related with the ith output and the nth power term. Eq. (12) could be phrased in a compact form as follows

$$\hat{y}_i^{MP}(t) = (\mathbf{z}_i(t))^T \boldsymbol{\alpha}_i, \tag{13}$$

where $\mathbf{z}_i(t) = [(z_i(t))^1; (z_i(t))^2; \ldots; (z_i(t))^N]$ and $\boldsymbol{\alpha}_i = [\boldsymbol{\alpha}_{1i}; \boldsymbol{\alpha}_{2i}; \ldots; \boldsymbol{\alpha}_{Ni}]$. To perform the identification of the MP model, let us consider the set of data $u(t)$ and $y_i(t)$ for $t = 1, 2, \ldots, K$, where K is the number of the available sampled data.

Now, it is possible to write

$$\hat{\mathbf{y}}_i = \mathbf{Z}_i \boldsymbol{\alpha}_i, \tag{14}$$

where $\mathbf{Z}_i = [(\mathbf{z}_i(1))^T; \ldots; (\mathbf{z}_i(K))^T]$ is the matrix of regression vectors, and $\hat{\mathbf{y}}_i = [\hat{y}_i^{MP}(1); \ldots; \hat{y}_i^{MP}(K)]$ is the vector of estimated outputs. Grouping the output data in a vector $\mathbf{y}_i = [y_i(1); \ldots; y_i(K)]$ it is possible to compute the parameter vector which are to be identified through least squares (LS), i.e., minimizing the squared error defined as

$$\hat{\boldsymbol{\alpha}}_i = \arg\min_{\alpha_i} (\mathbf{y}_i - \mathbf{Z}_i \boldsymbol{\alpha}_i)^T (\mathbf{y}_i - \mathbf{Z}_i \boldsymbol{\alpha}_i), \tag{15}$$

where the solution is given as

$$\hat{\boldsymbol{\alpha}}_i = \left(\mathbf{Z}_i^T \mathbf{Z}_i\right)^{-1} \mathbf{Z}_i^T \mathbf{y}_i. \tag{16}$$

In order to ensure the invertibility of the term $\mathbf{Z}_i^T \mathbf{Z}_i$, the condition of persistent excitation should be satisfied [2].

2.2. MGMP model

In this case, the model output is (Fig. 1(a))

$$\hat{y}_i^{MGMP}(t) = \sum_{n=1}^{N} (\boldsymbol{\alpha}_{ni}^T z_i(t))^n, \tag{17}$$

where $\boldsymbol{\alpha}_{ni} \in Re^{N_{z_i}}$ is the parameter vector related with the ith output and the nth power term. Since that the output is not linear in the parameters, it is not possible to use the classical LS solution similar to the one presented in Eq. (16). Then, the proposal is to solve the LS using an optimization program.

The approximation error for the output i at time t is given as

$$e_i(t) = y_i(t) - \sum_{n=1}^{N} (\boldsymbol{\alpha}_{ni}^T z_i(t))^n, \tag{18}$$

and the identification could be performed by solving the following problem

$$\hat{\boldsymbol{\alpha}}_i = \arg\min_{\alpha_i} \sum_{t=1}^{K} (e_i(t))^2. \tag{19}$$

Note that this is an unconstrained nonlinear optimization problem. In order to improve the convergence of this identification algorithm, both gradient and Hessian analytical expressions are provided (see Appendix A).

3. Identification of a distillation column

In this section, the model identification methodology presented above is illustrated with a simulation example. In this case, a distillation column has been selected as a case study.

Distillation columns comprise a significant portion of separation processes in process industries [12]. Because distillation operation directly affects product quality, process production rates and utility usage, the economic importance of distillation control is undisputed. Modeling and control of distillation process dynamics is a challenging problem due to process nonlinearity, multivariable coupling, severe disturbances and non-stationary behavior.

Therefore, in the areas of modeling and control, distillation columns have attracted the attention of several researchers. Such is the case of Skögestad and Morari [13] and Skögestad and Postlethwaite [14], whose Column A has been widely studied. This simulation example is herein selected to illustrate the use of the identification methodology.

In this case the LV control structure is used. The input $u = [V_B\ L_T]^T$ is a vector formed by the boilup and the reflux flows, respectively. On the other hand, the output $y = [x_B\ x_D]^T$ is a vector formed by the liquid bottom composition and the liquid distillate composition, respectively. Therefore, we deal with a two input-two output process for the identification.

Simulation of the nonlinear model was performed to collect the necessary input–output data of this nonlinear process. For such purpose, random signals with uniform distribution around 2% of the nominal steady-state operating point were considered for the inputs (i.e. manipulated variables). A sample time of 25 s was assumed and the input was maintained constant for 40 samples. A total amount of 8000 samples were obtained, half of them were used for identification and the rest for model validation.

The dominant poles in the Laguerre basis were chosen taking into account a preliminary linear identification. In this case, a Laguerre expansion of order 1 was selected with poles $\xi_{1,1} = 0.95,\ \xi_{1,2} = 0.95,\ \xi_{2,1} = 0.95$ and $\xi_{2,2} = 0.80$. The orthonormal basis is presented as a state space model as

$$z(t+1) = Az(t) + Bu(t),$$

$$z_i(t) = C_i z(t), \quad \text{for} \quad i = 1, 2,$$

where

$$A = \begin{bmatrix} 0.9500 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.0975 & 0.9500 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.9500 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.0975 & 0.9500 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.9500 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.0975 & 0.9500 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.8000 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.3600 & 0.8000 \end{bmatrix},$$

$$B = \begin{bmatrix} 0.3122 & 0 \\ -0.2966 & 0 \\ 0 & 0.3122 \\ 0 & -0.2966 \\ 0.3122 & 0 \\ -0.2966 & 0 \\ 0 & 0.6000 \\ 0 & -0.4800 \end{bmatrix},$$

$$C_1 = [I_4 \quad 0_4],$$

$$C_2 = [0_4 \quad I_4],$$

where I_4 is the identity matrix of dimension 4 and 0_4 is the zero matrix of dimension 4.

MP and MGMP models were identified based on the collected data. In order to have a comparison index, a linear model and a model based on Piecewise Linear Approximation (PWL) were also obtained [15]. This last one is known as a model with excellent approximation properties, then it is useful as a reference of model quality. As regards the polynomials, order 3 was selected, while for the PWL partition, each dimension was split into 5 segments [15]. Table 1 shows the necessary number of

parameters as well as the approximation and validation errors for both outputs, where the error determination was performed according to the following criterion

$$E = \sum_{i=1}^{2}\sum_{t=1}^{K}(e_i(t))^2. \tag{20}$$

Figs. 4 and 5 show both the model output and measured data when modeling with the MGMP structure. The two composition variables (i.e., x_B and x_D) are depicted. The MGMP structure approach was selected to illustrate the identification approach, specially due to the better approximation results achieved with this model.

Table 1
Approximation and validation errors for different approaches.

Model	E_{app}	E_{val}	Number of parameters
Linear	1.7947	2.5948	8
PWL	0.3860	0.6236	2592
MP	0.7298	1.3585	24
MGMP	0.3070	0.5861	24

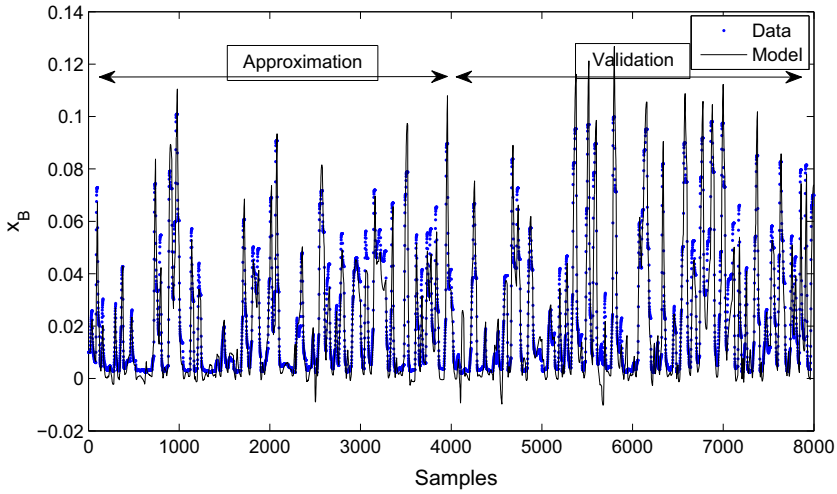

Fig. 4. Bottom composition (x_B): data and model prediction.

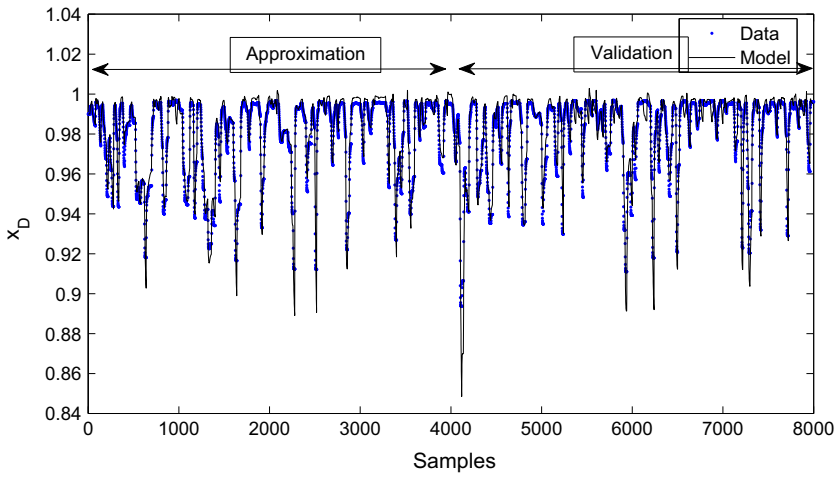

Fig. 5. Distillate composition (x_D): data and model prediction.

Table 2
Computational demand assessment.

Model	Stage I (s)	Stage II (s)
PWL	46.5	71.80
MP	0.1	0.2
MGMP	1.5	0.6

Table 2 gives details on the computational demand (evaluated as time consumption in seconds) associated with the compared identification methods. Stage I represents both data generation and training task (i.e., model parameters identification). Otherwise, Stage II stands for the simulation task. For Stage I 4000 input–output data were used, while for simulation purpose (i.e. Stage II), a set of 8000 data were considered.

The optimisation involved in the MGMP approach, demanded 75 iterations for the first output and 104 iterations for the second one.

4. Conclusions

A dedicated approach for identification of nonlinear systems was presented. A new proposal for the representation and the identification of MP and MGMP models (both Volterra-like models) was dealt with.

In a comparative study of both approaches (MP and MGMP structures), it was obtained that, provided the same amount of parameters is included in each model, different level of approximation is achieved. In the study case herein considered, MGMP significantly outperformed MP approximation capability. Furthermore, both of them clearly surpassed PWL model approximation index, with the additional advantage that they require far fewer parameters than the PWL model. From the identification results it emerges another observation in favor of MGMP model, which is the good coherence between approximation and validation errors.

As regards the identification procedure associated to each model structure (i.e., PWL, MP and MGMP representations), it should be noted that neither PWL nor MP use optimisation, both approaches involve matrix inversion instead. Otherwise, identification of MGMP model is formulated as an optimisation problem. This fact is reflected in the computational effort requirement. In this sense, computational assessment results show a substantial advantage as regards relative values for time demand. However, both "absolute" values are in the same order of magnitude (the second), hence the difference could be negligible for many subsequent applications (for instance, applications in process control).

Therefore, a global appreciation is both MP and MGMP models allow satisfactory approximation results for a reasonable amount of model parameters. Although MGMP model evidences better approximation performance, the selection between both models may be subject to the particular system to model and/or further involved applications.

Acknowledgments

The authors gratefully acknowledge the financial support of Agencia Nacional de Promoción Científica y Tecnológica, CONICET and Universidad Nacional del Sur.

Appendix A

To improve the convergence of the optimization problem, the computation of both gradient and Hessian analytical expressions is accomplished. The objective function of the problem for the ith output, is

$$f_i(\boldsymbol{\alpha}) = \sum_{t=1}^{K}\left(y_i(t) - \sum_{n=1}^{N}(\boldsymbol{\alpha}_{ni}^T z_i(t))^n\right)^2. \tag{21}$$

The gradient of $f_i(\boldsymbol{\alpha})$ with respect to the entries of $\boldsymbol{\alpha}_i$ is

$$\frac{\partial f_i(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}_{ni}} = -2n\sum_{t=1}^{K} e_i(t)(\boldsymbol{\alpha}_{ni}^T z_i(t))^{n-1} z_i(t) \tag{22}$$

and the Hessian is

$$\frac{\partial^2 f_i(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}_{ni}\partial \boldsymbol{\alpha}_{mi}} = -2n\sum_{t=1}^{K}\left(-m(\boldsymbol{\alpha}_{mi}^T z_i(t))^{m-1}(\boldsymbol{\alpha}_{ni}^T z_i(t))^{n-1} z_i^2(t) + \psi(n-1)e_i(t)(\boldsymbol{\alpha}_{ni}^T z_i(t))^{n-2} z_i^2(t)\right), \tag{23}$$

where $\psi = 1$ only if $m = n$; otherwise $\psi = 0$.

References

[1] P.H. Menold, F. Allgöwer, R.K. Pearson, Nonlinear structure identification of chemical processes, Comput. Chem. Eng. 21 (1997) S137–S142.
[2] M. Schetzen, The Volterra and Wiener Theories of Nonlinear Systems, John Wiley & Sons, USA, 1980.
[3] F.J. Doyle III, R.K. Pearson, Identification and Control Using Volterra Models, Springer, London, Great Britain, 2002.
[4] S. Boyd, L.O. Chua, Fading memory and the problem of approximating nonlinear operators with Volterra series, IEEE Trans. Circuits Syst. CAS-32 (11) (1985) 1150–1161.
[5] A. Janczak, Identification of Nonlinear Systems Using Neural Networks and Polynomial Models, Springer, Berlin, 2005.
[6] S. Biagiola, J.L. Figueroa, Wiener and Hammerstein uncertain models identification, Math. Comput. Simul. 79 (11) (2009) 3296–3313.
[7] D.R. Morgan, Z. Ma, J. Kim, M.G. Zierdt, J. Pastalan, A generalized memory polynomial model for digital predistortion of RF power amplifiers, IEEE Trans. Circuits Syst. 54 (10) (2006) 3852–3860.
[8] N. Björsell, Modeling analog to digital converters at radio frequency, Doctoral Thesis in Telecommunications Stockholm, Sweden, 2007.
[9] C. Schmidt, J.E. Cousseau, J.L. Figueroa, R. Wichman, S. Werner, Nonlinearities modeling and post-compensation in continuous-time sigma–delta modulators, IET Microwaves Antennas Propag. 5 (15) (2011) 1796–1804.
[10] B. Wahlberg, System identification using Kautz models, IEEE Trans. Autom. Control 39 (6) (1994) 1276–1282.
[11] B. Wahlberg, P.M. Mäkilä, On approximation of stable linear dynamic systems using Laguerre and Kautz functions, Automatica 32 (5) (1996) 693–708.
[12] J. Riggs, M. Karim, Chemical and Bio-Process Control, third ed., Pearson Education International, Boston, USA, 2007.
[13] S. Skögestad, M. Morari, Understanding the dynamic behaviour of distillation columns, Ind. Eng. Chem. Res. 27 (10) (1998) 1848–1862.
[14] S. Skögestad, I. Postlethwaite, Multivariable Feedback Control. Analysis and Design, second. ed., John Wiley & Sons, 2007.
[15] M.P. Alvarez, L.R. Castro, O.E. Agamennoni, Set membership estimation theory for Wiener modelling using HLCPWL functions, Int. J. Model. Ident. Control 14 (1–2) (2011) 13–26.