IEEE Instrumentation and Measurement
Technology Conference
Anchorage, AK, USA, 21-23 May 2002, pp. 887-892.

SVD Based Modeling of Nonlinear Systems

A.R. Várkonyi-Kóczy[1,2], P. Baranyi[2,3], A.E. Ruano[4], S. Győri[1,2], Z. Petres[1,2], P. Légrády[5]
[1]Dept. Measurement and Information Systems, Budapest University of Technology and Economics
Magyar Tudósok krt. 2., Budapest, H-1521 Hungary, e-mail: koczy@mit.bme.hu
[2]Integrated Intelligent Systems Japanese-Hungarian Laboratory
[3]Dept. of Telecommunication and Telematics, Budapest University of Technology and Economics
Magyar Tudósok krt. 2., Budapest, H-1521 Hungary, e-mail: baranyi@alpha.ttt.bme.hu
[4]Dept. of Electronic Engineering and Computing, University of Algarve
Campus de Gambelas, 8000 Faro, Portugal, email: aruano@ualg.pt
[5]Universmed, Budapest University of Technology and Economics
Műegyetem rkp. 1-3, Budapest, H-1521 Hungary

Abstract *– Nowadays practical solutions of measurement and control problems involve model-integrated computing. Model based approaches offer a very challenging way to integrate a priori knowledge into the procedure. In this paper the applicability of the Higher Order Singular Value Decomposition based modeling is presented together with a medical example, the modeling of the behavior of the human liver-bile system.*

Keywords *– nonlinear modeling, HOSVD, model approximation, product-sum gravity approximation, human liver-bile system*

I. INTRODUCTION

Nowadays practical solutions of measurement and control problems involve model-integrated computing. This integration means that the available knowledge is represented in a proper form and acts as an active component of the computer program to be executed during the operation of the measuring and control devices. Thus, model based approaches offer a very challenging way to integrate a priori knowledge into the procedure.

In case of linear problems very well established methods are available and they are successfully combined with adaptive techniques to provide optimum performance. Nonlinear techniques, however, are far from this maturity. There is a wide variety of possible models to be applied based on both classical methods [1] and recent advances in handling information [2] but up till now practically no systematic method was available which could be offered to solve a larger family of nonlinear engineering problems. The efforts on the fields of fuzzy and neural network (NN) modeling and control seem to result in a real breakthrough in this respect (see e.g. [3]). These techniques can be applied even in cases when no analytical knowledge is available about the system, the information is uncertain or inaccurate, or when the available mathematical form is too complex to be used. Although, the use of fuzzy and NN models is limited by their exponentially increasing complexity and we can easily get into trouble when the problem is very complex with probably large number of parameters especially when only partial and uncertain measured data is available about the system to be modeled.

Singular Value Decomposition (SVD) technique has successfully been used to reduce the complexity of a large family of systems based on both classical and soft techniques [4]. An important advantage of the SVD reduction technique is that it offers a formal measure to filter out the redundancy (exact reduction) and also the weakly contributing parts (non-exact reduction). This implies that the degree of reduction can be applied according to the maximum acceptable error corresponding to the temporal circumstances.

All these initiated the extension of the application of SVD not only for complexity reduction but for modeling, as well. The main idea behind is that if we have the analytical description of the system or if we have a system represented by input-output sample pairs, we can build a model approximating the system using local (linear) models. Such techniques include Takagi-Sugeno (TS) fuzzy model approximation [5] or polytopic model approximation (PMA). These methods have theoretically a universal approximation property, however, it can not be really exploited because they have an exponentially increasing complexity growing by the number of parameters. These means that if the number of the local units is bounded then the built model will only be an approximation of the original system. Thus, we have to find a balance between the computational complexity and the accuracy.

On the other hand, after ensuring a needed or given accuracy, which may mean the application of a huge number of local models, the computational complexity can be reduced by applying some kind of exact or non-exact complexity reduction method.

To get a universal multi-dimensional modeling method, the above technique can be defined in a tensor product form, i.e. Higher Order SVD (HOSVD) can be applied. It can also be extended with a progressively refining partitioning.

The advantages of the proposed method are the low computational complexity, the easy realization, that the accuracy can easily and flexibly increased, that we do not need any a priori or expert information about the system to be modeled, etc. It can be used either when a mathematical description (possibly too complex to be handled) or when only (possibly partial, inaccurate, and uncertain) measurement data is available. The technique can be applied in anytime systems, as well.

The paper is organized as follows: in Section II. As basis for the chosen model approximation method the singleton based Product-Sum Gravity (PSG) inference is summarized. Section III. is devoted to SVD and Higher Order SVD (HOSVD) complexity reduction techniques. In Section IV. a medical problem, the behavior of the human liver-bile system is outlined for which the modeling process is presented in Section V. Section VI. draws the Conclusions.

II. SINGLETON BASED PRODUCT-SUM-GRAVITY INFERENCE METHOD

Let assume a fuzzy rule base given by antecedent fuzzy sets $A_{in,n}$ defined on input universe X_n and by singleton consequent fuzzy sets b_m , on output universe Y. $1 \le n \le N$ denotes the input dimension, $1 \le i_n \le I_n$ the number of fuzzy sets in dimension n, and $1 \le m \le M$, where M is the number of fuzzy sets in Y.

The rules are

IF $A_{i1,1}$ AND $A_{i2,2}$ AND … AND $A_{iN,N}$ THEN $b_{i1,i2,\dots,iN}$.

During the inference product T-norm, sum S-norm, and center of gravity inference are used:

$$\hat{y}(x) = \sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \cdots \sum_{i_N=1}^{I_N} \prod_{n=1}^{N} \mu_{A i_n,n}(x_n) b_{i_1,i_2,\dots i_N}$$

$\hat{y}(x)$ $\mu_{Ain,n}$ indicates the membership functions, and is the output of the system. From this we can equivalently formulate a tensor product form expression

$$\hat{y} = B \otimes \boldsymbol{\mu}_n$$

Let also suppose that the fuzzy sets are in Ruspini partition:

$$\sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \cdots \sum_{i_N=1}^{I_N} \prod_{n=1}^{N} \mu_{i_n,n}(x) = 1.$$

With this assumptions the process of the modeling is the following: Over a rectangular hyper-grid the approximate model is built of local models and basis functions (fuzzy sets) serving for the superposition realizing the whole system. In the simplest case the local models are constant values ($b_{i1,i2,\dots,iN}$) set in the grid-points and the basis functions ($\mu_{Ain,n}$) are of triangular shape (linear approximation, TSMA) with such operational areas that the effect of the local models is limited by the basis functions to the areas bounded by the neighboring grid-points. In general the proposed technique is not limited to this case. Any function can serve as basis function (b-splines, trigonometric functions, etc.), the only limitation is that they have to be in Ruspini partition.

The technique realizes a Lagrange interpolation and the aim of the approximation is to copy the behavior of the unknown system in the given sample (input) points. If we are free to define the location of the samples (e.g. a differential equation system is given which can be sampled [4]), they can be located in the grid-points, which means that the samples can take the role of the local models making much easier the modeling since the effect of the local models (fuzzy sets) is limited for two "sections" around a grid-point bounded by two others. In this case in every grid-point one and only one fuzzy set/parameter will have effect.

If we work with pre-given sampling points which usually are not set in the grid-points, we have to adapt the model to approximate the behavior of the system in the sampling points and not in the grid points. Thus, the modeling means the determination of the elements of tensor B (i.e. the local models located in the grid points) supposing that the basis functions are fixed in advance.

We can determine the elements of tensor B with the help of the training patterns. The model fitting can be carried out by using some kind of adaptation technique, e.g. Least Mean Square (LMS) method based on the available measured data. To each input value the response of the model y is calculated. To minimize the output, the error criteria of the LMS algorithm is used. In the kth time point the error criteria ($C^{(k)}$) is

$$C^{(k)} = \left(E^{(k)}\right)^2 = \left(y^{(k)} - \hat{y}^{(k)}\right)^2 = \left(y^{(k)} - \sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \cdots \sum_{i_N=1}^{I_N} \prod_{n=1}^{N} \mu_{i_n,n}(x_n) b^{(k)}_{i_1,i_2,\dots i_N} \right)^2$$

The determination of the elementary models means the minimization of $C^{(k)}$ ($E^{(k)}$ is the error, $y^{(k)}$ denotes the output of the system to be modeled), by the well known iterative procedure

$$b^{(k+1)}_{i_1,i_2,\ldots,i_N} = b^{(k)}_{i_1,i_2,\ldots,i_N} + \eta\left(-\nabla^{(k)}_{i_1,i_2,\ldots,i_N}\right) =$$
$$= b^{(k)}_{i_1,i_2,\ldots,i_N} + 2\eta E^{(k)} \mu_{i_1,1}(x_1)\mu_{i_2,2}(x_2)\cdots\mu_{i_N,N}(x_N).$$

$\nabla^{(k)}_{i_1,i_2,\ldots,i_N}$ stands for the derivative of the error criteria according to the unknown elemental model, η is the learning factor.

In higher order cases, i.e. when $N>2$ the modeling procedure can also be expressed in a tensor form.

An important quest is the density of the approximation points. It can be proved that the tensor product interpolation method is universal approximator. On the other hand the complexity of the system grows exponentially with the accuracy. The more points and bases are taken into consideration the more improved model we get. Moreover, using a too fine hyper-grid means that such operational domains will be defined in the model about which we possibly do not have any information. It causes that the output surface will be not smooth enough. The areas where we do not have any measurement point will remain "untouched" by the algorithm. This fact suggests a method where we step by step re-fine the grid and interpolate the intermediate points.

First a rough partition is chosen and the training procedure is applied on this surface. As a next step, the partitioning is refined in such a way that the new grid points are placed in center of the sections defined by the old grid points. After the refinement process, the number of points increases from I_n(old) to I_n(new)=2 I_n(old)-1 in each dimension and tensor B will also be modified according to the new points. If the approximation error drops to a certain level, i.e. the number of points becomes acceptable in a dimension no more refinement will be needed in this dimension.

III. HIGHER ORDER SVD (HOSVD) BASED COMPLEXITY REDUCTION

Fuzzy logic approximators suffer from exponentially growing computational complexity in respect to their approximation property. Furthermore, in most of the cases the number of fuzzy sets are overestimated because of the lack of any systematic method which would be suitable for estimating the necessary number of sets to ensure a given accuracy. Thus, after finishing the model-building phase the computational complexity is to be optimized by applying e.g. the SVD complexity reduction technique [6]. For this we should limit the applied basis functions for those that can be expressed as rational fractions.

The aim of the complexity reduction is twofold. First, we would like to filter out the redundancy of the model (exact reduction), second, if necessary we can further decrease the computational complexity by applying non-exact reduction. In anytime systems this latter can be down adaptively according to the temporal conditions [7], [8].

The SVD-based complexity reduction algorithm is based on the decomposition of the $F_{(n1xn2)}$ matrix

$$F = A_1 D A_2^T = \begin{bmatrix} A_1^r & A_1^d \end{bmatrix} \begin{bmatrix} B^r & 0 \\ 0 & B^d \end{bmatrix} \begin{bmatrix} A_2^r & A_2^d \end{bmatrix}^T.$$

Matrix D is diagonal containing the singular values in decreasing order. The zero or the smallest of singular values (smaller than singular value threshold T_0, say) can be discarded from diagonal matrix D to yield a simpler system. Matrix B^r thus contains the retained singular values and B^d contains the discarded ones. The result of SVDR is therefore

$$\hat{F} = A_1^r B^r {A_2^r}^T = A_1^r B'.$$

For n-dimension cases Higher Order SVD (HOSVD) can be applied similarly (see e.g. [9]).

The error caused by non-exact reduction can be upper-bounded in cases of both the basic SVD and HOSVD by the sum of the discarded singular values. If SVD is applied then the reduction will be optimal, i.e. it results a minimum system. If HOSVD is applied then the situation is a bit different. The minimum property does not hold anymore we can state only that the "significant" singular values will have lower indexes, however in most cases it results in an approximation which is very near to the possible best one.

In our case HOSVD is applied on tensor B, thus after the reduction the number of local models will be (nearly) minimized.

The method also offers a way for improving the model if later on we get into possession of new approximation points. An algorithm can be suggested which finds the common minimal implementation space of the compressed original and the new approximation points (for more details see [10]).

IV. THE HUMAN LIVER-BILE SYSTEM

The human liver-bile system might be most unknown and most complex part of the human organism. It seems so that there is a close relation between the changes of the different blood components, hormone levels and the functioning of the human liver-bile system. Until recently, medical researchers mostly examined only small parts instead of the whole complex system and, as the measurement results show, the resulting models seem to be incorrect. A typical example of this is the front-page story of the Time Magazine [11] saying that the consumption of eggs is unhealthy for people and increases significantly the level of cholesterol. 15 years later, in another issue [12], another article appeared stating exactly the opposite also in a cover story.

The liver is our system's chemical factory, it produces substances of vital importance that influence the functioning of the entire body. The base material of the factory is the consumed nutriment, thus some malfunctioning of the digestive system can be traced back to the unsuitable nutrition [13].

There are direct and indirect methods about this system. Direct methods are e.g. the measurement of cholesterol, bile acids, and hormone levels, but this information is not enough to get to know the whole system. The indirect methods not only give more detailed information but they are much cheaper. In our analysis the following measurement methods were used to tune the model. After a substantial breakfast, the emptying process of the gall bladder is measured by ultrasound: by taking two perpendicular cuts of the bladder the volume is calculated from the length of different dimensions as an ellipsoid. This is repeated every 20-30 minutes. The level of cholesterol, triglyceride, and bile acids in the blood are measured. The following information is available for every patient: initial values of the three blood components, the initial volume of the gall bladder (for empty stomach), the volume of the gall bladder after a substantial breakfast in 20-30-minute periods (these data points define the function $g(t)$, the undefined values of $g(t)$ are calculated by linear approximation), and on the morning of the next day, values of the three blood components, and the morning volume of the gall bladder. Our examination is based only on the cholesterol values from the blood components. Let us indicate with C and C' the initial and the next morning cholesterol values.

In the analysis of the functioning of the gall bladder, not the values of $g(t)$ but rather its shape, its characteristics is important. For this reason, four new parameters are determined that describe the important features of the behavior of the function $g(t)$:

$V_i = g(0)$ (initial volume),
$V_m = \min g(t)$ (minimal volume),
t_e when $g(t_b)=V_m + (V_k-V_m)/10$ first (time when the volume enters the 10% stripe),
t_l when $g(t_k)= V_m + (V_k-V_m)/10$ last (time when the volume leaves the 10% stripe during a measured time period T (see Fig.1).

Our goal was to create a model based on five input variables t_e, t_l, V_i, V_m, C to be able to reproduce C' as output.

The system is highly nonlinear, uncertain and inaccurate thus the robust soft computational method presented in the previous sections was used.

V. THE MODEL

In this Section the numerical results of the model of the human liver-bile system are presented. The domains of the five input parameters were partitioned uniformly in each dimension. The size of the (refined) final tensor B is 9x9x9x9x9. The initial values of the grid-points were uniformly chosen to 0.

The antecedent fuzzy sets were half overlapping triangular functions. The value of the learning rate η was between 0.1 and 0.001. In each partition refinement step tensor B was trained by 5000 inputs, and to ensure the needed accuracy the number of grid points were doubled in 5 steps in each dimension (some steps of the refinement process are shown in Fig.2).

Since the size of tensor B is large, HOSVD complexity reduction was applied. During the singular value decompositions, those singular values were dropped which proved to be smaller then 0.05. The size of the reduced obtained tensor B^r is 8x7x7x9x9, which is only 53.8% of the original one. The measured maximal L_2 error of the reduced model is $2.9128 \cdot 10^{-3}$, thus the obtained model is suitable for further medical analysis.

VI. CONCLUSIONS

Nowadays, model based approaches have become very popular. Although, there is a lack of such systematic methods which can be applied for a larger class of nonlinear problems, as well. In this paper a Higher Order Singular Value Decomposition based universal modeling approach is presented which can be advantageously used in case of highly nonlinear problems, where only partial, uncertain and/or inaccurate data is available. The usability of the method is illustrated by a medical example, where the behavior of the human liver-bile system is modeled.

REFERENCES

[1] S.A. Billings, "Identification of Nonlinear Systems – A Survey," *IEE Proc.* Vol. 127, Pt.D, No. 6, Nov. 1980., pp. 272-284.

[2] G.J. Klir, T.A. Folger, *Fuzzy Sets, Uncertainty, and Information*, Prentice-Hall International, Inc., 1988.
[3] R. Babuska, J. Sousa, H.B. Verbruggen, "Model-Based Design of Fuzzy Control Systems," *Proc. of the Third European Congress on Intelligent Techniques and Soft Computing EUFIT'95*, Aachen, Germany, pp. 837-841.
[4] Y. Yam, "Fuzzy Approximation via Grid Sampling and Singular Value Decomposition," *IEEE Trans. S.M.C.*, Vol. 27, No. 6, 1997, pp. 933-951.
[5] T. Takagi, and M. Sugeno, "Fuzzy Identification of Systemsand its Applications to Modeling and Control," *IEEE Trans. S.M.C.*, Vol. 15, 1985, pp. 116-132.
[6] Y. Yam, P. Baranyi, C.T. Yang, "Reduction of Fuzzy Rule Base Via Singular Value Decomposition," *IEEE Trans. on Fuzzy Systems*, Vol. 7, No. 2, 1999, pp. 120-131.
[7] A.R. Várkonyi-Kóczy, A. Ruano, P. Baranyi, O.Takács, "Anytime Information Processsing Based on Fuzzy and Neural Network Models," *Proc. of the 2001 IEEE Instrumentation and Measurement Technology Conference, IMTC/2001*, Budapest, Hungary, May 21-23, 2001, pp. 1247-1252.
[8] O.Takács, A.R. Várkonyi-Kóczy, "Iterative-type Evaluation of Fuzzy Systems for Anytime Use," *Proc. of the 2002 IEEE Instrumentation and Measurement Technology Conference, IMTC/2002,* Anchorage, USA, May 21-23, 2002.
[9] Baranyi, P., A.R. Várkonyi-Kóczy, Y.Yam, P.Michelberger, "HOSVD Based Computational Complexity Reduction of TS Fuzzy Models," *Proc. of the Joint 9th IFSA World Congress and 20th NAFIPS International Conference, IFSA / NAFIPS 2001*, Vancouver, Canada, July 25-28, 2001, pp.
[10] P. Baranyi, A.R. Várkonyi-Kóczy, "Adaptation of SVD Based Fuzzy Reduction via Minimal Expansion," *Proc. of the IEEE Int. Workshop on Intelligent Signal Processing, WISP'2001*, Budapest, Hungary, May 24-25, 2001, pp. 171-176.Reference
[11] "Cholesterol. And Now the Bad News…" *Time Magazine*, March 26, 1984.
[12] "Cholesterol. And Now the Good News…" *Time Magazine*, July 19, 1999.
[13] P. Légrády, *Egg, Nutrition, Health*, Maecenas Könyvkiadó, Budapest, 2001, (In Hungarian).

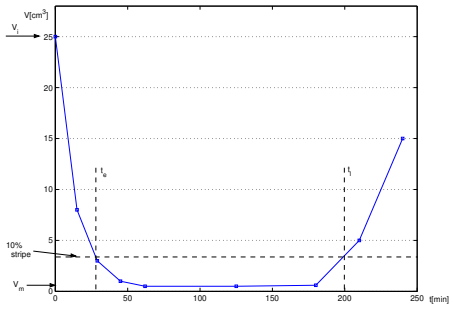

Fig.1 A typical function ($g(t)$) for the operation of the gall bladder

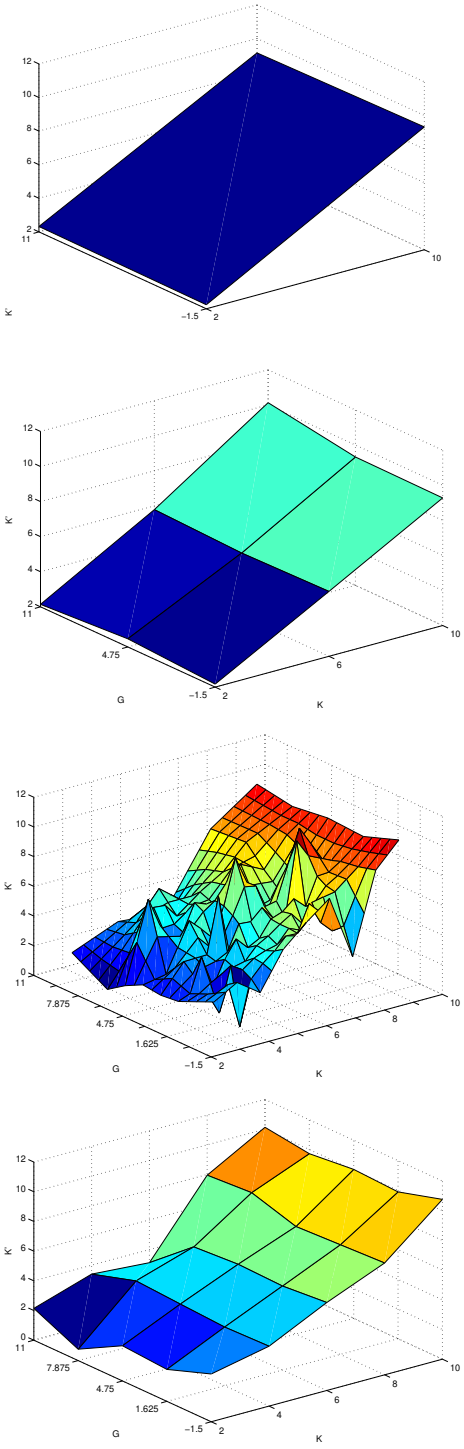

Fig.2 The refinement process