Dynamics Modeling of Multidimensional Interrelated Processes

Serhiy Yefimenko

International Research and Training Centre of Information Technologies and Systems of the NAS and MES of Ukraine, Akademik Glushkov prospect 40, 03680 Kyiv, Ukraine

syefim@ukr.net

Abstract. *The paper considers theoretical grounds of recurrent computing applying for dynamics modeling of multidimensional interrelated processes. The combinatorial GMDH algorithm is used for vector autoregressive (VAR) modelling. The procedure of structural and parametric identification of VAR models is proposed.*

Keywords

Inductive modelling, VAR, recurrent parameters estimation, combinatorial GMDH algorithm

5 Introduction

An approach to the problem of structural and parametric identification of vector autoregressive models, when parameters for every model are estimated independently, is considered. The drawback of the approach is that the parameters of separate models for interrelated processes are interdependent. The paper uses an algorithm with selection not one but a few best models for every process to eliminate the shortcoming.

We use the Group Method of Data Handling (GMDH) [1] as one of the most effective inductive modelling methods. When solving the problem of vector autoregressive modelling, the run-time is one of the most important criteria of efficiency. The most effective way to achieve the high performance is recurrent parameter estimation. In this case it is more expedient to use the high-speed methods of parameters estimation based on the recurrent algorithms for solving of linear equations systems.

2 Vector Autoregressive Modelling

The vector autoregressive (VAR) model generalizes the model of autoregression to multidimensional case. It is built by the stationary time series. It is the system of equations, in which every variable (component of multidimensional time series) is linear combination of all variables in the previous time points. The order of such model is determined by the order of the lags.

In the general case for m time series and k lags the model will be the system of m equations (in matrix form):

$$X(t) = \sum_{j=1}^{k} \Theta_j X(t-j), \tag{1}$$

where Θ_j, $j = \overline{1,k}$ – matrices of model parameters (3) of size $m \times m$.

3 Procedure of Structural and Parametric Identification of VAR Models

The general models structure in the form of a system of m difference equations is determined as a result of the sequence

of such operations:

1. Data array of $m \cdot k$ arguments is composed under the number of interrelated processes m and lags κ.
2. Maximal complexity for restricted search is defined. COMBI algorithm with sequentially complicated structures of models on the basis of recurrent computing is used for modelling. For every time series the best F (by the value of the regularity criterion [2]) models are selected. Overall $F \cdot m$ models are passed to the next step.
3. The sorting of $G = F^m$ possibilities of systems of models is carried out. The best model (by the value of the systemic integral criterion of vector models quality) is selected. The value of the criterion is calculated on the given part of initial data set in the prediction mode of the process for the given steps number n_S:

$$B = \sum_{i=1}^{n_S} \sum_{j=1}^{m} (x_{ij} - x_{ij}^*)^2 , \tag{2}$$

where x_{ij}^* is the result of step-by-step integration (recursive computations) of the system of m equations.

4 Combinatorial GMDH Algorithm with Recurrent Parameters Estimation

The combinatorial algorithm is used for VAR modelling by exhaustive search of all possible variants and finding the best model for every time series containing the most informative subset of input arguments (regressors). It consists of such main units:

- data conversion according to a basic class of models (linear in parameters);
- forming models of different complexity;
- calculation values of external quality criteria for all models being formed;
- selection of the best models.

For linear object with m input, all possible models are compared in the process of exhaustive search. Total quantity of all generated models of the type

$$\hat{y}_v = X_v \hat{\theta}_v , \; v = 1, \ldots, 2^m - 1 \tag{3}$$

is 2^m–1. Decimal number v corresponds to binary number d_v in (3). Unit elements of d_v indicate inclusion regressors with corresponding numbers in the model, whereas zero elements signify exclusion.

Due to the exponential growth of 2^m as function of arguments amount, it is advisable to use algorithms recurrent in the number of parameters in structural identification problems for the parameters estimation.Efficient recurrent modifications of classic Gauss and Gramm-Schmidt algorithms were offered in [3]. The modification, in a nutshell, is as follows. The matrix $H_s = X_s^T X_s$ of the size $s \times s$ is reduced to superdiagonal form by computing only elements h_{is}^s, $i = \overline{2, s-1}$, h_{si}^s, $i = \overline{2, s}$, and $g_s = X_s^T y$ at every step s, $s = \overline{1, m}$ during the direct motion. The elements of the nested matrix H_{s-1} of size $(s-1) \times (s-1)$ (reduced to superdiagonal form on the previous step) remain changeless. So only "bordering elements" are computed on step s.

5 Conclusion

The technique of recurrent computing in COMBI GMDH algorithm for building discrete prediction models of complex multidimensional interrelated processes is proposed. Software tools for modelling of complex multidimensional interrelated processes on the basis of high-performance combinatorial algorithm in the class of discrete dynamic VAR models are developed.

References

[5] Madala, H.R., Ivakhnenko, A.G.: Inductive Learning Algorithms for Complex Systems Modeling. CRC Press, New York (1994).
[6] Farlow, S.J.: Self-Organizing Methods in Modelling: GMDH Type Algorithms. New-York, Bazel: Marcel Decker Inc., (1984)
[7] Stepashko, V.S., Efimenko S.N.: Sequential Estimation of the Parameters of Regression Model. Cybernetics and Systems Analysis 41(4), 631–634 (2005).