УДК 681.5.015

EXPERIMENTAL VERIFICATION OF INTERNAL CONVERGENCE OF ITERATIVE GMDH ALGORITHMS

Volodymyr Stepashko[1], Oleksandra Bulgakova[2], Viacheslav Zosimov[1]

1. *International Research and Training Centre for Information Technologies and Systems of the NAS and MES of Ukraine. Akademik Glushkov Prospect 40, Kyiv, 03680, Ukraine*
2. *Mykolaiv V.O.Suhomlynsky National University*

sashabulgakova1@gmail.com; stepashko@irtc.org.ua

В статті за допомогою обчислювальних експериментів перевіряється наявність так званої внутрішньої збіжності деяких ітераційних алгоритмів МГУА.
Ключові слова: індуктивне моделювання, МГУА, багаторядний ітераційний алгоритм, комбінаторний алгоритм, гібридний алгоритм

The availability of so-called internal convergence of some iterative GMDH algorithms is verified in this paper using computational experiments.
Keywords Inductive modeling, GMDH, multilayered iterative algorithm, combinatorial algorithm, hybrid algorithm

В статті за допомогою обчислювальних експериментів перевіряється наявність так званої внутрішньої збіжності деяких ітераційних алгоритмів МГУА.
Ключевые слова: индуктивное моделирование, МГУА, многорядный итерационный алгоритм, комбинаторный алгоритм, гибридный алгоритм

Iterative algorithms

In the classical multilayered iterative algorithm MIA GMDH [1] each partial model $f(x_i, x_j)$ is formed as a combination of input variables x_i, x_j, $i, j = 1,2,...,m$, in initial layer and combination of model outputs obtained in the previous layer, starting from the second. Typically linear, bilinear or quadratic polynomials of a pair of arguments may be used as the partial descriptions:

$$y_l^r = a_0 + a_1 y_i^{r-1} + a_2 y_j^{r-1}, \quad (1)$$

$$y_l^r = a_0 + a_1 y_i^{r-1} + a_2 y_j^{r-1} + a_4 y_i^{r-1} y_j^{r-1}, \quad (1a)$$

$$y_l^r = a_0 + a_1 y_i^{r-1} + a_2 y_j^{r-1} + a_3 (y_i^{r-1})^2 + a_4 y_i^{r-1} y_j^{r-1} + a_5 (y_j^{r-1})^2 \quad (1b)$$

Coefficients $a_0, a_1, \ldots, a_5$ are estimated in the training data set using the least squares method. The best *F* models with respect to the value of an external selection criterion are involved in the formation of model variants in a next layer. The process is finished in the layer, after which the criterion minimum begins to grow. In this case there is a possibility of essential arguments loss in the selection process because those lost after the first layer fall out from the subsequent selection process.

In [2] we proposed the following types of hybrid algorithms:

1. Combined Iterative Algorithm (CIA) – the algorithm with equal usage of both intermediate and initial arguments:

$$y_l^{(r+1)} = f(y_i^r, y_j^r) \vee f(y_i^r, x_j) \quad (2)$$

2. Combined Iterative Combinatorial Algorithm (CICA) – algorithm with combinatorial optimization of partial descriptions complexity and adding the initial arguments in each layer to avoid loosing the relevant ones

$$y_l^{(r+1)} = f_{opt}(y_i^r, y_j^r) \vee f_{opt}(y_i^r, x_j) \quad (3)$$

Combinatorial optimization consists in that on each layer we consider the model, for example linear:

$$f(y_i^r, y_j^r) = a_0 d_1 + a_1 d_2 y_i^{r-1} + a_2 d_3 y_j^{r-1}, \quad (4)$$

where $d_k, k = 1,2,3$ are elements of the binary structural vector d that takes on a value 1 or 0 (inclusion or non-inclusion of the proper argument): $d_k = \{0, 1\}$, $f_{opt}(y_i^r, y_j^r) = f(y_i^r, y_j^r, d_{opt})$:

$$\left.\begin{array}{ll}
100 \to f_1 = a_0 & CR_1 \\
010 \to f_2 = \quad a_1 y_i^{r-1} & CR_2 \\
001 \to f_3 = \quad\quad a_1 y_j^{r-1} & CR_3 \\
110 \to f_4 = a_0 + a_1 y_i^{r-1} & CR_4 \\
101 \to f_5 = a_0 + \quad + a_1 y_j^{r-1} & CR_5 \\
011 \to f_6 = \quad a_1 y_i^{r-1} + a_2 y_j^{r-1} & CR_6 \\
111 \to f_7 = a_0 + a_1 y_i^{r-1} + a_2 y_j^{r-1} & CR_7
\end{array}\right\} \underset{\min}{\Rightarrow} f_{opt}$$

Internal convergence of iterative algorithms

An iterative algorithm holds an internal convergence if as a result of iteration process with the selection criterion *CR*=*RSS*, where *RSS* is the residual sum of squares on the entire sample *W* (without partition), the estimation of optimal model parameters are converging to LSM estimates of the true structure model parameters.

The sample splitting to the training and testing sub-samples is not used in this case (m=10, n=50). It is studying internal convergence of the three iterative GMDH algorithms: classical MIA, combined CIA, and generalized CICA.

The true model is the following linear function of ten arguments:

$$\overset{\circ}{y} = 3 - 2x_1 + 5x_2 - x_3 + 7x_4 - 2x_5 + 4x_6 - 3x_7 + 2{,}5x_8 - 1{,}8x_9 + x_{10}. \quad (5)$$

Table 1 and Fig. 1 shows the change of RSS values as dependent on the layer number r of these algorithms with $F = 20$.

Table 1

RSS value in the layers for the three iterative algorithms

Layer, r	RSS for CIA	RSS for CICA	RSS for MIA
1	416,15	416,15	416,15
2	194,25	121,75	346,52
3	44,02	9,20	309,34
4	5,10	0,98	238,12
5	1,47	0,53	201,7
6	0,70	0,27	144,78
7	0,65	0,11	113,12
8	0,59	0,003	113,06
9	0,36	0,0011	–
10	0,21	0,0010	–
11	0,13	0,0010	–
12	0,007	–	–
13	0,0069	–	–
14	0,0068	–	–
15	0,0066	–	–
16	0,0064	–	–

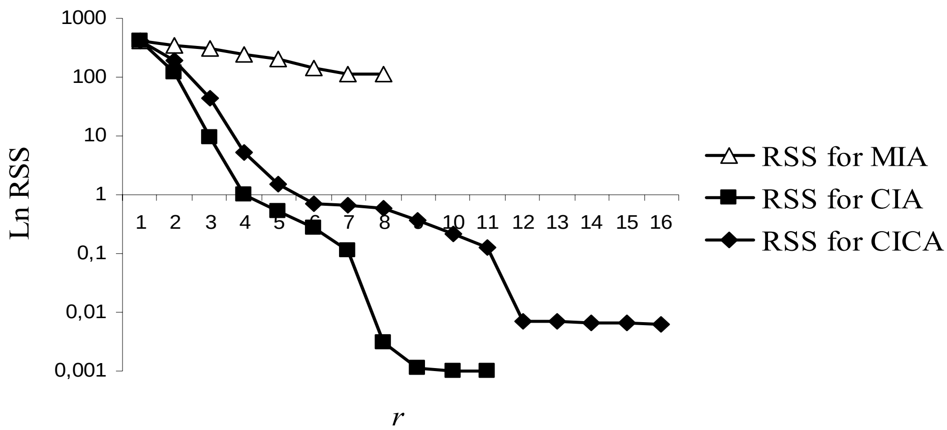

Fig. 1 Change of RSS values for the three iterative GMDH algorithms, $F = 20$ (logarithmic scale)

The constructed models are of form:

$$y_{\text{CIA}} = 2{,}913 - 1{,}978x_1 + 4{,}999x_2 - 1{,}015x_3 + 7{,}004x_4 - 1{,}988x_5 + \\ + 3{,}997x_6 - 2{,}994x_7 + 2{,}514x_8 - 1{,}808x_9 + 1{,}003x_{10} \quad (6)$$

$$y_{\text{CICA}} = 3{,}001 - 1{,}999x_1 + 4{,}999x_2 - 1{,}002x_3 + 7{,}004x_4 - 1{,}998x_5 + \\ + 4{,}000x_6 - 2{,}999x_7 + 2{,}500x_8 - 1{,}801x_9 + 1{,}003x_{10} \quad (7)$$

$$y_{\text{MIA}} = -22,387 + 5,334x_2 - 0,535x_3 + 5,266x_4 + 0,812x_5 + 4,726x_6 + 1,278x_8 - 2,643x_9 + 0,437x_{10} \quad (8)$$

The first and second equations (6), (7) generally confirm the internal convergence of the two iterative algorithms as values of the estimated parameters are close to the true ones, see (5). The fig.1 shows that the generalized algorithm reaches its minimum on the 9th layer and the combined one on the 12th, so the convergence rate of CICA is much higher than CIA. The accuracy of the obtained models also differs in favor of the generalized algorithm: for CICA *RSS*=0.001, for CIA *RSS*=0.006. This demonstrates the effectiveness of the idea of partial models optimization. In the classical multilayered algorithm MIA the convergence is observed only by criterion but there is no internal convergence with respect to the structure and parameters in view of the loss of true arguments x_1 and x_7, see (8).

Table 2 shows the change in the best models structure and parameters for CICA. One can see that the true structure was found at 4 layer after which there was only changing the values of the parameters being estimated.

Table 2

RSS value in the best models structure and parameters for CICA

Layer r	RSS	Model
1	416,15	$\hat{y} = -15,806 + 5,521x_2 + 6,581x_4$
2	121,75	$\hat{y} = -7,220 + 4,434x_2 + 5,286x_4 + 5,604x_6 - 1,988x_9$
3	9,20	$\hat{y} = 7,084 + 4,782x_2 + 5,700x_4 - 1,073x_5 + 4,467x_6 - 3,412x_7 + 2,177x_8 - 2,647x_9 + 0,905x_{10}$
4	0,98	$\hat{y} = 2,245 - 1,953x_1 + 5,028x_2 - 1,019x_3 + 7,054x_4 - 2,008x_5 + 4,016x_6 - 2,968x_7 + 2,492x_8 - 1,801x_9 + 0,962x_{10}$
5	0,53	$\hat{y} = 3,121 - 1,982x_1 + 4,987x_2 - 1,021x_3 + 6,986x_4 - 1,989x_5 + 4,016x_6 - 2,968x_7 + 2,505x_8 - 1,801x_9 + 0,964x_{10}$
6	0,27	$\hat{y} = 3,121 - 1,982x_1 + 4,987x_2 - 1,015x_3 + 6,986x_4 - 1,995x_5 + 4,004x_6 - 2,995x_7 + 2,505x_8 - 1,801x_9 + 0,965x_{10}$
7	0,11	$\hat{y} = 3,011 - 1,989x_1 + 4,997x_2 - 1,003x_3 + 7,005x_4 - 1,998x_5 + 4,001x_6 - 2,997x_7 + 2,501x_8 - 1,801x_9 + 0,988x_{10}$
8	0,003	$\hat{y} = 3,001 - 1,999x_1 + 4,999x_2 - 1,003x_3 + 7,005x_4 - 1,998x_5 + 4,000x_6 - 2,999x_7 + 2,501x_8 - 1,801x_9 + 1,003x_{10}$
9	0,0011	$\hat{y} = 3,001 - 1,999x_1 + 4,999x_2 - 1,002x_3 + 7,004x_4 - 1,998x_5 + 4,000x_6 - 2,999x_7 + 2,501x_8 - 1,801x_9 + 1,003x_{10}$
10	0,0010	$\hat{y} = 3,001 - 1,999x_1 + 4,999x_2 - 1,002x_3 + 7,004x_4 - 1,998x_5 + 4,000x_6 - 2,999x_7 + 2,500x_8 - 1,801x_9 + 1,003x_{10}$

To confirm this results, 10 redundant arguments was added to the data sample, hence m = 20, and true linear model corresponds to the same formula (5). When doing so, only Combined Iterative Combinatorial algorithm was used. The following models were obtained using LSM (regression analysis) and CICA:

$$\begin{aligned}\hat{y}_{REGR} &= 3-2x_1+4{,}998x_2-x_3+6{,}999x_4-2x_5+3{,}999x_6-\\&-3x_7+2{,}498x_8-1{,}8x_9+1{,}000x_{10}-0{,}000001x_{11}+\\&+0{,}000002x_{12}-0{,}00001x_{13}+0{,}0000003x_{14}+0{,}0000001x_{15}+\\&+0{,}00001x_{16}-0{,}0000003x_{17}+0{,}000001x_{18}+0{,}000002x_{19}+0{,}00001x_{20}\end{aligned} \quad (9)$$

$$\begin{aligned}\hat{y}_{CICA} &= 3{,}001-1{,}999x_1+4{,}999x_2-1{,}002x_3+7{,}004x_4-1{,}998x_5+4{,}000x_6-\\&-2{,}999x_7+2{,}501x_8-1{,}801x_9+1{,}003x_{10}-0{,}00001x_{13}+0{,}000003x_{19}\end{aligned} \quad (10)$$

The model (9) built by LSM includes all 20 arguments, both true and redundant, though with very small coefficients, and RSS = 0.0001 for it.

The model (10) confirms the CICA algorithm effectiveness even with 10 redundant (uninformative) arguments. This model includes redundant arguments x_{13} and x_{19} but with very small values of estimated parameters.

This means that the iterative algorithm CICA GMDH can be used not only for structural and parametric identification but also as an iterative procedure of parameter estimation of models by minimizing RSS.

Conclusion

The results obtained above illustrate, first, that the developed architecture of the Combined Iterative Combinatorial GMDH algorithm allows investigating various versions of iterative algorithms, and second, this algorithm demonstrates higher performance in testing tasks.

The complex technique of analyzing the efficiency of iterative GMDH algorithms using computational experiments allows a comparative study of various iterative algorithms effectiveness.

The comparison of the hybrid CICA GMDH algorithm with other iterative GMDH algorithms has been made based on study of the internal convergence process. Executed experiments showed the best performance of the CICA algorithm.

References

[1] Ivakhnenko A.G. Group method of data handling as a competitor for the method of stochastic approximation. – Soviet Automatic Control, 1968, No. 3. – P. 58-72.

[2] Stepashko V., Bulgakova O., Zosimov V. Modified multilayered GMDH algorithm with combinatorial optimization of partial descriptions complexity // Proceedings of the III Intern. Conf. on Inductive Modelling ICIM-2010, May 16-22, 2010, Yevpatoria, Crimea, Ukraine. – Kherson: KNTU, 2010. – P. 24-29.