

А.П. Сарычев

**АЛГОРИТМ ПОИСКА ОПТИМАЛЬНОГО МНОЖЕСТВА РЕГРЕССОРОВ В ЗАДАЧЕ ВЕКТОРНОЙ РЕГРЕССИИ**

*Аннотация. В соответствии с принципами метода группового учета аргументов построен многоэтапный итерационный алгоритм структурной идентификации, позволяющий находить модель оптимальной сложности в задаче векторной регрессии в условиях структурной неопределённости.*

*Ключевые слова: метод группового учёта аргументов, критерий регулярности.*

При решении задач одномерной и векторной регрессии, поставленных в условиях структурной неопределённости, как правило, предполагается, что генерация наборов регрессоров выполняется полным перебором всех возможных сочетаний исходного множества регрессоров. Полный перебор реализуется в виде многоэтапных процедур, в которых на каждом этапе с номером  $s$  в модели допускается не более  $s$  регрессоров. Генерация наборов регрессоров начинается с отдельно взятых регрессоров на этапе с номером  $s=1$  и заканчивается на этапе  $s^*$ , на котором попытка усложнить модель по числу включенных в нее регрессоров не приводит к улучшению качества решения. Для многих практических приложений задача полного перебора регрессоров на современных ЭВМ является вполне реализуемой: об-

щее число перебираемых вариантов равно  $\sum_{s=1}^{s^*} C_m^s$ , где  $C_m^s = \frac{m! s!}{(m-s)!}$  –

число сочетаний из  $m$  элементов по  $s$ . Но существуют практические задачи, поставленные таким образом, что число регрессоров в исходном множестве  $X$  слишком велико. В этих условиях алгоритм полного перебора заменяют эвристическими процедурами, которые реализуют так называемые "рациональные", "целенаправленные" способы перебора. Безусловно, этапу применения таких эвристических проце-

дур должно предшествовать их исследование методом статистических испытаний.

В методе группового учёта аргументов (МГУА) [1–2] существуют два класса алгоритмов поиска оптимальной структуры регрессии: комбинаторные [1] и итерационные алгоритмы [2–4]. Один из представителей второго класса – МИАСИС – многоэтапный итерационный алгоритм структурной идентификации (системный).

### 1. Назначение алгоритма

В случае одномерной по выходу регрессии многоэтапный итерационный алгоритм структурной идентификации (МИАСИ) [5–6] позволяет определить структуру и оценить коэффициенты зависимости скалярной выходной переменной от входных переменных в условиях, когда априорно неизвестно, какие именно входные переменные участвуют в её формировании и какова дисперсия её случайной составляющей. Зависимости могут быть линейными по коэффициентам, линейными или нелинейными по входным переменным. Генерация, сравнение и отбор моделей проводятся по многоэтапной схеме, в которой номер этапа определяет максимально возможное число членов в моделях. На первом этапе анализируются все возможные одночлены, на втором – двухчлены. На последующих этапах структура и сложность генерируемых моделей зависит от структуры лучших моделей предыдущего этапа. Для оценивания качества моделей выборка данных разбивается на две подвыборки: на обучающей подвыборке по итерационной схеме МГУА оцениваются коэффициенты модели, а на проверочной подвыборке нормой остатков оценивается её качество. Оптимальная (лучшая по структуре) модель определяется по минимуму нормы остатков на проверочной подвыборке (критерий регулярности). Предусмотрено, что качество моделей может оцениваться и по схеме скользящего экзамена (усреднённый критерий регулярности), а также остаточной суммой квадратов на обучающей подвыборке.

Класс моделей, синтезируемых по алгоритму МИАСИС, в случае одномерной по выходу регрессии имеет вид

$$y = \sum_{g=1}^s \theta_g \cdot \prod_{j=1}^m x_j^{\alpha(g,j)}, \quad (1)$$

где  $y$  – выходная переменная;  $g$  – номер члена в модели,  $g = 1, 2, \dots, s$ ;  $s$  – число членов в модели;  $\theta_g$  – коэффициент при  $g$ -ом члене;  $x_j$  –

$j$ -я входная переменная,  $j = 1, 2, \dots, m$ ;  $m$  – число входных переменных;  $\alpha(g, j)$  – показатель степени, в которой  $j$ -я входная переменная входит в  $g$ -й член модели.

Класс моделей (1) в частном случае, когда выходная переменная линейно зависит от входных переменных, имеет вид:

$$y = \sum_{g \in Q(p,s)} \theta_g \cdot x_g, \quad s = 1, 2, \dots, m; \quad p = 1, 2, \dots, P; \quad P = C_m^s, \quad (2)$$

где  $Q(p, s)$  – одно из всевозможных подмножеств множества  $\{1, 2, \dots, m\}$  такое, что число элементов в нём равно  $s$ ;  $P$  – число таких всевозможных подмножеств, равное числу сочетаний из  $m$  элементов по  $s$ .

Имея наблюдения выходной и входных переменных

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad \mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix} \quad (3)$$

и предполагая существование зависимости выходной переменной от набора входных переменных в указанных классах моделей (1)–(2), можно по программе алгоритма МИАСИ [5–6] определить структуру модели (параметры  $\alpha(g, j)$  и  $s$ ) и оценить её коэффициенты  $\theta_g$ ,  $g = 1, 2, \dots, s$ .

## 2. Многоэтапный итерационный алгоритм структурной идентификации

Алгоритм МИАСИС позволяет определить структуру и оценить коэффициенты зависимости многомерной (векторной) выходной переменной от входных переменных в условиях, когда априорно неизвестно, какие именно входные переменные участвуют в формировании каждой выходной переменной, и неизвестна ковариационная матрица их случайных составляющих. Если коэффициенты регрессионных моделей для каждой из компонент векторной выходной переменной можно оценивать независимо, т. е. отдельно для каждой скалярной входной переменной, то поиск оптимальных по структуре регрессионных моделей можно проводить по алгоритму МИАСИ, применяя его отдельно для каждой компоненты векторной выходной переменной. В случае, когда коэффициенты регрессионных моделей

для всех компонент векторной выходной переменной необходимо оценивать совместно, следует применять системный аналог алгоритма МИАСИ – алгоритм МИАСИС. Класс моделей, синтезируемых по алгоритму МИАСИС в системе регрессионных уравнений, совпадает с (1)–(2). Имея векторы наблюдений  $h$  выходных переменных и соответствующие матрицы наблюдений входных переменных ( $k = 1, 2, \dots, h$ ):

$$y(k) = (y_1(k), y_2(k), \dots, y_n(k))^T, \quad X(k) = \begin{bmatrix} x_{11}(k) & x_{12}(k) & \cdots & x_{1,m(k)}(k) \\ x_{21}(k) & x_{22}(k) & \cdots & x_{2,m(k)}(k) \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1}(k) & x_{n2}(k) & \cdots & x_{n,m(k)}(k) \end{bmatrix}, \quad (4)$$

и, предполагая существование зависимостей выходных переменных от набора входных переменных в указанных классах моделей (1)–(2), можно по программе алгоритма МИАСИС определить структуры моделей (параметры  $\alpha(k; g, j)$ ,  $s(k)$ ) и оценить их коэффициенты  $\theta_g(k)$ ,  $g = 1, 2, \dots, s(k)$ .

Разобьем исходную выборку наблюдений на обучающую ( $A$ ) и проверочную ( $B$ ) подвыборки; сгруппируем векторы наблюдений выходных переменных  $y(k)$ ,  $k = 1, 2, \dots, h$ , и матрицы наблюдений входных переменных  $X(k)$ ,  $k = 1, 2, \dots, h$ , следующим образом:

$$y(k) = \begin{pmatrix} y(A, k) \\ y(B, k) \end{pmatrix}, \quad X(k) = \begin{bmatrix} X(A, k) \\ X(B, k) \end{bmatrix}, \quad (5)$$

где  $y(A, k)$ ,  $X(A, k)$  –  $(n(A) \times 1)$ -вектор наблюдений выходной переменной с номером  $k$  и  $(n(A) \times m(k))$ -матрица наблюдений её входных переменных на обучающей подвыборке;  $y(B, k)$ ,  $X(B, k)$  –  $(n(B) \times 1)$ -вектор наблюдений выходной переменной с номером  $k$  и  $(n(B) \times m(k))$ -матрица наблюдений её входных переменных на проверочной подвыборке;  $n(A)$  и  $n(B)$  – объемы обучающей и проверочной подвыборок,  $n(A) + n(B) = n$ .

Дадим необходимые определения:

1) частным описанием для выходной переменной с номером  $k$  будем называть  $(n \times 1)$ -вектор  $\hat{z}(k)$ , полученный по алгоритму МИАСИС как некоторое приближение (аппроксимация) вектора  $y(k)$ ;

2) *структурой* частного описания для выходной переменной с номером  $k$  будем называть набор параметров  $\alpha(k; g, j)$  и  $s(k)$ , определяющих  $\hat{z}(k)$  в представлении (1).

Прежде чем приступить к построению алгоритма, определим общий вид матрицы частных описаний для выходной переменной с номером  $k$ :

$$\begin{aligned} \hat{\mathbf{Z}}^r(k) &= \left[ \begin{array}{c|c|c|c} \hat{\mathbf{z}}_1^r(k) & \hat{\mathbf{z}}_2^r(k) & \dots & \hat{\mathbf{z}}_{F+2+m+2 \times s^F(k)}^r(k) \end{array} \right] = \\ &= \left[ \begin{array}{c|c|c|c} \hat{\mathbf{z}}_1^r(A, k) & \hat{\mathbf{z}}_2^r(A, k) & \dots & \hat{\mathbf{z}}_{F+2+m+2 \times s^F(k)}^r(A, k) \\ \hline \hat{\mathbf{z}}_1^r(B, k) & \hat{\mathbf{z}}_2^r(B, k) & \dots & \hat{\mathbf{z}}_{F+2+m+2 \times s^F(k)}^r(B, k) \end{array} \right] = \left[ \begin{array}{c} \hat{\mathbf{Z}}^r(A, k) \\ \hline \hat{\mathbf{Z}}^r(B, k) \end{array} \right], \end{aligned} \quad (6)$$

где  $r=1, 2, \dots$  – номер итерации;  $\hat{\mathbf{z}}_j^r(k)$  –  $(n \times 1)$ -векторы, частные описания;  $j=1, 2, \dots, F+2+m+2 \times s^F$ ;  $n$  – число наблюдений;  $F$  – число лучших частных описаний, передаваемых от итерации к итерации (и от этапа к этапу);  $m(k)$  – число входных переменных для  $k$ -й выходной переменной;  $s^F(k)$  – число членов в структуре лучшего частного описания  $(r-1)$ -й итерации для выходной переменной с номером  $k=1, 2, \dots, h$ .

Для удобства дальнейшего изложения осуществим разбиение:

$$\hat{\mathbf{Z}}^r(k) = \left[ \begin{array}{c|c|c} \hat{\mathbf{G}}^r(k) & \hat{\mathbf{C}}^r(k) & \hat{\mathbf{D}}^r(k) \end{array} \right], \quad (7)$$

$$\hat{\mathbf{G}}^r(k) = \left[ \begin{array}{c|c|c|c} \hat{\mathbf{z}}_1^r(k) & \hat{\mathbf{z}}_2^r(k) & \dots & \hat{\mathbf{z}}_F^r(k) \end{array} \right], \quad (8)$$

$$\hat{\mathbf{C}}^r(k) = \left[ \begin{array}{c|c|c|c} \hat{\mathbf{z}}_{F+1}^r(k) & \hat{\mathbf{z}}_{F+2}^r(k) & \dots & \hat{\mathbf{z}}_{F+2+m(k)}^r(k) \end{array} \right], \quad (9)$$

$$\hat{\mathbf{D}}^r(k) = \left[ \begin{array}{c|c|c|c} \hat{\mathbf{z}}_{F+2+m(k)+1}^r(k) & \hat{\mathbf{z}}_{F+2+m(k)+2}^r(k) & \dots & \hat{\mathbf{z}}_{F+2+m(k)+2 \times s^F(k)}^r(k) \end{array} \right]. \quad (10)$$

Для построения итерационного алгоритма МГУА необходимо [3, 4]:

1) указать начальные матрицы частных описаний  $\hat{\mathbf{Z}}^0(k)$ ,  $k = 1, 2, \dots, h$ ;

2) определить оператор  $\mathfrak{R}$ , осуществляющий отображение

$$\left[ \hat{\mathbf{Z}}^{r-1}(1) \mid \hat{\mathbf{Z}}^{r-1}(2) \mid \dots \mid \hat{\mathbf{Z}}^{r-1}(h) \right] \rightarrow \left[ \hat{\mathbf{Z}}^r(1) \mid \hat{\mathbf{Z}}^r(2) \mid \dots \mid \hat{\mathbf{Z}}^r(h) \right], \quad (11)$$

где  $r = 1, 2, \dots$  – номер итерации;

3) указать правило завершения итераций.

Шаг 1. Укажем начальную матрицу частных описаний  $\hat{\mathbf{Z}}^0(k)$  (положим пока  $s^F(k) = 0$ ):

$$\hat{\mathbf{Z}}^0(k) = \left[ \mathbf{O}_{n \times F} \mid \mathbf{0}_n \mid \mathbf{1}_n \mid \mathbf{X}(k) \right], \quad k = 1, 2, \dots, h, \quad (12)$$

где  $\mathbf{O}_{n \times F}$  – нулевая  $(n \times F)$ -матрица;  $\mathbf{0}_n$  – нулевой  $(n \times 1)$ -вектор;  $\mathbf{1}_n$  – единичный  $(n \times 1)$ -вектор;

$$\mathbf{X}(k) = \left[ \mathbf{x}_1(k) \mid \mathbf{x}_2(k) \mid \dots \mid \mathbf{x}_{m(k)}(k) \right] - \quad (13)$$

$(n \times m(k))$ -матрица  $n$  наблюдений  $m(k)$  входных переменных для выходной переменной с номером  $k$ .

Шаг 2. Определим оператор  $\mathfrak{R}$ .

Пусть  $(n \times 1)$ -векторы  $\hat{\mathbf{z}}^r(k)$ ,  $k = 1, 2, \dots, h$ , определяются по правилу:

$$\hat{\mathbf{z}}_i^r(k) = \hat{a}(A, k) \cdot \hat{\mathbf{z}}_{i, j_1}^{r-1}(k) + \hat{b}(A, k) \cdot \hat{\mathbf{z}}_{i, j_2}^{r-1}(k) \cdot \hat{\mathbf{z}}_{i, j_3}^{r-1}(k), \quad (14)$$

где  $r = 1, 2, \dots$  – номер итерации;  $i = 1, 2, \dots, n$  – номер наблюдения;

$j_1, j_2, j_3 = 1, 2, \dots, F + 2 + m(k) + 2 \times s^F(k)$  ( $j_2 \leq j_3$ ) – номера частных опи-

саний из матрицы  $\hat{\mathbf{Z}}^{r-1}(k)$ ;  $\hat{a}(k)$ ,  $\hat{b}(k)$  – коэффициенты, определяемые

на обучающей подвыборке наблюдений ( $A$ ) обобщённым методом наименьших квадратов как решение задачи параметрической идентификации системы регрессионных уравнений [7] – совместной регрессии выходных переменных  $y(k)$ ,  $k = 1, 2, \dots, h$ , причём каждая по двум своим регрессорам:

$$\hat{\mathbf{z}}_{j_1}^{r-1}(A, k) \quad \text{и} \quad \hat{\mathbf{z}}_{j_2}^{r-1}(A, k) \otimes \hat{\mathbf{z}}_{j_3}^{r-1}(A, k) \quad (15)$$

(здесь  $\otimes$  – знак покомпонентного перемножения векторов).

Для полученной регрессии выполняется

$$\mathbf{y}(A, k) = \hat{a}(k) \cdot \hat{\mathbf{z}}_{j_1}^{r-1}(A, k) + \hat{b}(k) \cdot \hat{\mathbf{z}}_{j_2}^{r-1}(A, k) \otimes \hat{\mathbf{z}}_{j_3}^{r-1}(A, k) + \mathbf{u}(A, k), \quad (16)$$

где  $(n \times 1)$ -вектор  $\mathbf{u}(A, k)$  представляет собой вектор остатков.

Из всех генерируемых по правилу (14)–(16) частных описаний отберём  $F$  описаний, имеющих наименьшие значения системного критерия регулярности  $ARS$  (исследован в [7]), который рассчитывается по остаткам регрессионных уравнений на проверочной выборке наблюдений ( $B$ ):

$$ARS = \frac{1}{h} \ln \left( \det \left[ \mathbf{U}^T(B | A) \mathbf{U}(B | A) \right] \right), \quad (17)$$

где

$$\mathbf{U}(B | A) = [\mathbf{u}(B | A, 1), \mathbf{u}(B | A, 2), \dots, \mathbf{u}(B | A, h)], \quad (18)$$

$$\mathbf{u}(B | A, k) = \mathbf{y}(B, k) - \hat{\mathbf{z}}^r(B, k), \quad k = 1, 2, \dots, h, \quad (19)$$

$$u_i(B | A, k) = y_i(B, k) - \hat{z}_i^r(B, k), \quad i = 1, 2, \dots, n(B), \quad (20)$$

$$\hat{z}_i^r(B, k) = \hat{a}(A, k) \cdot \hat{z}_{i, j_1}^{r-1}(B, k) + \hat{b}(A, k) \cdot \hat{z}_{i, j_2}^{r-1}(B, k) \cdot \hat{z}_{i, j_3}^{r-1}(B, k). \quad (21)$$

Эти наборы  $F$  лучших частных описаний, упорядоченные по возрастанию

системного критерия регулярности так, что выполняется

$$\begin{aligned} &ARS \left( \left[ \begin{array}{c|c|c|c} \hat{\mathbf{z}}_1^r(1) & \hat{\mathbf{z}}_1^r(2) & \dots & \hat{\mathbf{z}}_1^r(h) \end{array} \right] \right) \geq \\ &\geq ARS \left( \left[ \begin{array}{c|c|c|c} \hat{\mathbf{z}}_2^r(1) & \hat{\mathbf{z}}_2^r(2) & \dots & \hat{\mathbf{z}}_2^r(h) \end{array} \right] \right) \geq \dots \\ &\dots \geq ARS \left( \left[ \begin{array}{c|c|c|c} \hat{\mathbf{z}}_F^r(1) & \hat{\mathbf{z}}_F^r(2) & \dots & \hat{\mathbf{z}}_F^r(h) \end{array} \right] \right), \end{aligned} \quad (22)$$

образуют матрицы

$$\hat{\mathbf{G}}^r(k) = \left[ \begin{array}{c|c|c|c} \hat{\mathbf{z}}_1^r(k) & \hat{\mathbf{z}}_2^r(k) & \dots & \hat{\mathbf{z}}_F^r(k) \end{array} \right], \quad k = 1, 2, \dots, h. \quad (23)$$

Матрицы  $\hat{\mathbf{C}}^r(k)$  оставим без изменения:

$$\hat{\mathbf{C}}^r(k) = \hat{\mathbf{C}}^{r-1}(k), \quad k = 1, 2, \dots, h. \quad (24)$$

Матрицы  $\hat{\mathbf{D}}^r(k)$ ,  $k=1, 2, \dots, h$ , сформируем с учётом структуры лучшего частного описания  $\hat{\mathbf{z}}_F^r(k)$  из  $F$  отобранных частных описаний.

Первые  $s^F(k)$  столбцов матрицы  $\hat{\mathbf{D}}^r(k)$ ,  $k=1, 2, \dots, h$ , сформируем отдельными членами лучшего частного описания  $\hat{\mathbf{z}}_F^r(k)$ , выделяя их по правилу:

$$\hat{d}_{i,l}^r(k) = \hat{z}_{i,F+2+m(k)+l}^r(k) = \theta_l(k) \cdot \prod_{j=1}^m x_{ij}^{\alpha(k;l,j)}(k), \quad (25)$$

где  $l=1, 2, \dots, s^F(k)$  – номер члена в структуре;  $s^F(k)$  – число членов в структуре лучшего частного описания для  $k$ -й выходной переменной.

Вторые  $s^F(k)$  столбцов матрицы  $\hat{\mathbf{D}}^r(k)$  тоже сформируем на основе структуры лучшего частного описания  $\hat{\mathbf{z}}_F^r(k)$ , по очереди исключая из него отдельные члены по правилу ( $l=1, 2, \dots, s^F(k)$ ):

$$\hat{d}_{i,s^F(k)+l}^r(k) = \hat{z}_{i,F+2+m(k)+s^F(k)+l}^r(k) = \sum_{\substack{g=1 \\ (g \neq l)}}^s \theta_g(k) \cdot \prod_{j=1}^m x_{ij}^{\alpha(k;g,j)}(k). \quad (26)$$

Отображение

$$\left[ \hat{\mathbf{z}}^{r-1}(1) \mid \hat{\mathbf{z}}^{r-1}(2) \mid \dots \mid \hat{\mathbf{z}}^{r-1}(h) \right] \rightarrow \left[ \hat{\mathbf{z}}^r(1) \mid \hat{\mathbf{z}}^r(2) \mid \dots \mid \hat{\mathbf{z}}^r(h) \right], \quad (27)$$

осуществлено, оператор  $\mathfrak{R}$  определён.

Шаг 3. Укажем правило остановки для итерационной процедуры: вычисления заканчиваются на итерации  $r^*$ , если выполнено одно из условий

$$\begin{aligned} &ARS \left( \left[ \hat{\mathbf{z}}_F^{r^*}(1) \mid \hat{\mathbf{z}}_F^{r^*}(2) \mid \dots \mid \hat{\mathbf{z}}_F^{r^*}(h) \right] \right) = \\ &= ARS \left( \left[ \hat{\mathbf{z}}_F^{r^*+1}(1) \mid \hat{\mathbf{z}}_F^{r^*+1}(2) \mid \dots \mid \hat{\mathbf{z}}_F^{r^*+1}(h) \right] \right), \end{aligned} \quad (28)$$

$$\begin{aligned}
& ARS \left( \left[ \begin{array}{c|c|c|c} \hat{\mathbf{z}}_{F,r^*-1}^*(1) & \hat{\mathbf{z}}_{F,r^*-1}^*(2) & \dots & \hat{\mathbf{z}}_{F,r^*-1}^*(h) \end{array} \right] \right) - \\
& - ARS \left( \left[ \begin{array}{c|c|c|c} \hat{\mathbf{z}}_{F,r^*}^*(1) & \hat{\mathbf{z}}_{F,r^*}^*(2) & \dots & \hat{\mathbf{z}}_{F,r^*}^*(h) \end{array} \right] \right) \leq \delta_1,
\end{aligned} \tag{29}$$

где  $\delta_1 > 0$  – заданная величина.

Итерационную процедуру (6)–(29), проведенную для фиксированного  $s$ , будем называть *этапом* алгоритма МИАСИС. На  $s$ -ом этапе алгоритма не рассматриваются модели, у которых число членов больше  $s$ . Если информации о числе членов в регрессионной модели нет, то поиск проводится с этапа с номером  $s_1(k) = 1$  или с любого заданного  $s_1(k) > 1$ ,  $k = 1, 2, \dots, h$ .

Начальные матрицы частных описаний  $\hat{\mathbf{Z}}_0^{s_1(k)}(k)$ ,  $k = 1, 2, \dots, h$ , для начального этапа совпадают с (9), а для этапа  $s$  задаются по правилу

$$\hat{\mathbf{Z}}_0^s(k) = \hat{\mathbf{Z}}_{r^*}^{s-1}(k), \quad k = 1, 2, \dots, h, \tag{30}$$

где  $\hat{\mathbf{Z}}_{r^*}^{s-1}(k)$  – матрица лучших частных описаний, полученных на предыдущем этапе для  $k$ -й выходной переменной.

Вычисления заканчиваются на этапе с номером  $s^*$ , если выполнено одно из условий, аналогичных (28)–(29):

$$\begin{aligned}
& ARS \left( \left[ \begin{array}{c|c|c|c} \hat{\mathbf{z}}_{F,s^*}^*(1) & \hat{\mathbf{z}}_{F,s^*}^*(2) & \dots & \hat{\mathbf{z}}_{F,s^*}^*(h) \end{array} \right] \right) = \\
& = ARS \left( \left[ \begin{array}{c|c|c|c} \hat{\mathbf{z}}_{F,s^*+1}^*(1) & \hat{\mathbf{z}}_{F,s^*+1}^*(2) & \dots & \hat{\mathbf{z}}_{F,s^*+1}^*(h) \end{array} \right] \right),
\end{aligned} \tag{31}$$

$$\begin{aligned}
& ARS \left( \left[ \begin{array}{c|c|c|c} \hat{\mathbf{z}}_{F,s^*-1}^*(1) & \hat{\mathbf{z}}_{F,s^*-1}^*(2) & \dots & \hat{\mathbf{z}}_{F,s^*-1}^*(h) \end{array} \right] \right) - \\
& - ARS \left( \left[ \begin{array}{c|c|c|c} \hat{\mathbf{z}}_{F,s^*}^*(1) & \hat{\mathbf{z}}_{F,s^*}^*(2) & \dots & \hat{\mathbf{z}}_{F,s^*}^*(h) \end{array} \right] \right) \leq \delta_2,
\end{aligned} \tag{32}$$

где  $\hat{\mathbf{z}}_{F,s}^*(k)$  – лучшее частное описание этапа  $s$  (лучшее частное описание итерации  $r^*$  этого этапа) для  $k$ -й выходной переменной;  $\delta_2 > 0$  – заданная величина.

Алгоритм МИАСИ исследован аналитически в условиях повторных наблюдений [8, 9].

### 3. Заключение

В соответствии с принципами метода группового учета аргументов построен многоэтапный итерационный алгоритм структурной идентификации, позволяющий находить структуру оптимальной сложности в задаче векторной регрессии в условиях структурной неопределённости. Отличительными особенностями алгоритма являются: многоэтапный поиск модели; итерационное оценивание коэффициентов моделей; существование «свободы выбора» – от итерации к итерации и от этапа к этапу передаётся не одна, а заданное количество лучших моделей  $F$ ; применение внешнего критерия для сравнения и отбора моделей; возможность поиска оптимальной модели как в классе линейных, так и в классе нелинейных по входным переменным моделей; приёмы выделения отдельных членов лучшего частного описания и на этой основе расширение базисного множества аргументов; схема расчёта усреднённого критерия регулярности (критерия скользящего экзамена), оптимальная по вычислительным затратам; возможность оценивания коэффициентов как методом наименьших квадратов, так и методом наименьших модулей. Примеры применения итерационных алгоритмов МГУА для решения практических задач опубликованы в работах [10–14].

### ЛИТЕРАТУРА

1. Ивахненко А. Г. Помехоустойчивость моделирования / А.Г. Ивахненко, В. С. Степашко. – Киев : Наукова думка, 1985. – 216 с.
2. Ивахненко А. Г. Моделирование сложных систем по экспериментальным данным / А. Г. Ивахненко, Ю. П. Юрачковский. – М. : Радио и связь, 1987. – 120 с.
3. Юрачковский Ю. П. Сходимость многорядных алгоритмов МГУА / Ю. П. Юрачковский // Автоматика, 1981. – № 3. – С. 32–36.
4. Белозерский Е. А. Об одном подходе к построению многорядных алгоритмов МГУА с линейными частными описаниями / Е. А. Белозерский, Н. А. Ивахненко, Ю. П. Юрачковский // Автоматика, 1981. – № 5. – С. 3–7.
5. Сарычев А. П. Многорядный нелинейный гармонический алгоритм МГУА для самоорганизации прогнозирующих моделей / А. П. Сарычев // Автоматика. – 1984. – № 4. – С. 88–93.

6. Ивахненко А. Г. Самоорганизация динамических моделей с периодическими коэффициентами (на примере прогнозирования солнечной активности) / А. Г. Ивахненко, А. П. Сарычев // Автоматика. – 1985. – № 1. – С. 31–36.
7. Сарычев А. П. Моделирование в классе систем регрессионных уравнений на основе метода группового учета аргументов / А. П. Сарычев // Международный научно-технический журнал “Проблемы управления и информатики”. – 2013. – № 2. – С. 8–24.
8. Сарычев А. П. Решение проблемы разбиения в МГУА при расчете критерия регулярности в условиях активного эксперимента / А. П. Сарычев // Автоматика. – 1989. – № 4. – С. 19–27.
9. Сарычев А. П. Определение  $J$ -оптимального множества регрессоров по повторным выборкам наблюдений / А. П. Сарычев // Автоматика. – 1993. – № 3. – С. 58–66.
10. Сарычева Л. В. Компьютерный эколого-социально-экономический мониторинг регионов. Математическое обеспечение. / Л. В. Сарычева. – Днепропетровск : НГУ, 2003. – 222 с.
11. Sarycheva Ljudmila. Using GMDH for Modeling Economical Indices of Mine Opening / Ljudmila Sarycheva // System Analysis and Modeling Simulation (SAMS), Taylor & Francis. – 2003. – Vol. 43. – No. 10. – P. 1341–1349.
12. Алпатов А. П. Статистическое моделирование зависимостей энергетических возможностей и стоимости носителей от их технических характеристик / А. П. Алпатов, В. И. Кузнецов, А. П. Сарычев // Системные технологии. – Днепропетровск : ДНВП “Системные технологии”. – 2007. – № 1 (48). – С. 166–175.
13. Сарычева Л. В. Кластерно-регрессионный анализ финансовых показателей банков Украины на основе МГУА / Л. В. Сарычева, А. П. Сарычев // Індуктивне моделювання складних систем. – 2013. – Вип. 5. – С. 270–277.
14. Sarycheva L. GMDH-Clustering / L. Sarycheva, A. Sarychev // GMDH-Methodology and Implementation in C / Editor Godfrey Onwubolu / London : Imperial College Press, 2015. – P. 157–203.