

А.П. Сарычев

Линейная авторегрессия со случайными коэффициентами на основе метода группового учета аргументов в условиях квазиповторных наблюдений

Предложен критерий регулярности с разбиением выборок наблюдений на обучающие и проверочные подвыборки в условиях квазиповторных наблюдений для моделирования в классе авторегрессионных уравнений со случайными коэффициентами. Выявлено условие редукции оптимальной авторегрессионной модели, которое зависит от параметров модели и объемов выборок.

Ключевые слова: структурная неопределенность, критерий регулярности.

Запропоновано критерій регулярності з розбиттям вибірок спостережень на навчальні та перевіріні підвибіркі за умов квазіповторних спостережень для моделювання в класі авторегресійних рівнянь з випадковими коефіцієнтами. Виявлено умову редукції оптимальної авторегресійної моделі, що залежить від параметрів моделі і обсягів вибірок.

Ключові слова: структурна невизначеність, критерій регулярності.

Введение. Класс авторегрессионных уравнений со случайными коэффициентами – известный класс моделей в задачах структурной идентификации. Определение порядка авторегрессионных моделей в условиях неопределенности по количеству и составу входных переменных – актуальная задача в теории идентификации и управления, и для ее решения существуют различные подходы [1–8].

Эта задача есть одним из объектов исследования в методе группового учета аргументов (МГУА) [9–16], разработанного академиком НАН Украины Алексеем Григорьевичем Иващенко. Подход основан на разбиении выборки наблюдений на обучающую и проверочную части: на обучающей выборке оцениваются коэффициенты модели, а на проверочной – качество модели. В соответствии с принципами моделирования в МГУА, для того чтобы построить авторегрессионное уравнение оптимальной сложности, необходимо:

- указать метод оценивания коэффициентов в авторегрессионном уравнении;
- задать алгоритм генерирования авторегрессионных уравнений (структур моделей);
- разработать внешний критерий для оценки качества перебираемых структур;
- исследовать поведение математического ожидания внешнего критерия в зависимости от состава регрессоров;

- доказать существование авторегрессионной модели оптимальной сложности.

При моделировании в классе авторегрессионных уравнений в МГУА традиционно применяется критерий регулярности: обучающая выборка формируется первыми $n(A)$ наблюдениями временного ряда, проверочная выборка – последующими $n(B)$ наблюдениями, причем выполняется $n(A) + n(B) = n$, где n – объем исходного временного ряда. Применение критерия регулярности в таком виде нуждается в контроле, поскольку динамические свойства объекта могут проявляться неодинаково в разных фазах переходных процессов.

В данной статье предложено рассчитывать критерий регулярности в так называемой схеме квазиповторных наблюдений, возможной в условиях активного эксперимента. В этой схеме обучающая (A) и проверочная (B) выборки получены особым способом как пара реализаций функционирования объекта с близкими начальными условиями, качественно одинаковым характером переходных процессов и близкими состояниями в конечные моменты времени.

Априорные предположения о динамической системе

Пусть функционирование динамического объекта подчиняется закону в виде авторегрессионного уравнения со случайными коэффициентами

$$x_i = \begin{pmatrix} * & * & * \\ x_{i-1}, x_{i-2}, \dots, x_{i-p} \end{pmatrix} \begin{pmatrix} \theta_1(i-1) \\ \theta_2(i-2) \\ \vdots \\ \theta_p(i-p) \end{pmatrix} + \quad (1)$$

$$+ \zeta_{i-1} = \mathbf{Z}_{i,\bullet}^*(p) \Theta_{\bullet,i}(p) + \zeta_{i-1},$$

где x_i – ненаблюдаемое значение выходной переменной объекта в дискретные моменты времени $t = t_i$, $i = 1, 2, \dots, n$; n – общее число наблюдений; p – число предыдущих значений выходной переменной, влияющих на ее текущее значение; ζ_{i-1} – ненаблюдаемая случайная величина.

В модели функционирования (1) матрица $\mathbf{Z}(p)$ – $(n \times p)$ -матрица p предыдущих ненаблюдаемых значений переменной; в обозначении этой матрицы p означает, что в формировании величины x_i участвуют величины $x_{i-1}, x_{i-2}, \dots, x_{i-p}$:

$$\mathbf{Z}(p) = \begin{pmatrix} * & * & \dots & * \\ x_0 & x_{-1} & \dots & x_{1-p} \\ * & * & \dots & * \\ x_1 & x_0 & \dots & x_{2-p} \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \dots & * \\ x_{i-1} & x_{i-2} & \dots & x_{i-p} \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \dots & * \\ x_{n-1} & x_{n-2} & \dots & x_{n-p} \end{pmatrix} = \begin{pmatrix} * \\ \mathbf{Z}_{1,\bullet}^*(p) \\ * \\ \mathbf{Z}_{2,\bullet}^*(p) \\ \vdots \\ * \\ \mathbf{Z}_{i,\bullet}^*(p) \\ \vdots \\ * \\ \mathbf{Z}_{n,\bullet}^*(p) \end{pmatrix}, \quad (2)$$

а $\mathbf{Z}_{i,\bullet}^*(p)$ означает ту строку $\mathbf{Z}(p)$, в которой первый элемент – x_{i-1} .

В модели (1) матрица $\Theta(p)$ – $(p \times n)$ -матрица n реализаций $(p \times 1)$ -вектора ненаблюдаемых случайных величин; в обозначении этой матрицы p означает, что при формировании величины x_i i -я $(p \times 1)$ -строка матрицы $\mathbf{Z}(p)$ умножается на i -й $(p \times 1)$ -столбец матрицы $\Theta_{\bullet,i}(p)$.

Пусть относительно $(p \times 1)$ -вектора случайных коэффициентов $\Theta_{\bullet,i}(p)$ выполняется ($i = 1, 2, \dots, n$):

$$\Theta_{\bullet,i}(p) = \overset{\circ}{\theta} + \mathbf{H}_{\bullet,i}(p) = \overset{\circ}{\theta} + \eta_i(p), \quad (3)$$

где

$$\overset{\circ}{\theta} = (\overset{\circ}{\theta}_1, \overset{\circ}{\theta}_2, \dots, \overset{\circ}{\theta}_p)^T \quad (4)$$

– $(p \times 1)$ -вектор неизвестных детерминированных коэффициентов; $\mathbf{H}(p)$ – $(p \times n)$ -матрица n реализаций p ненаблюдаемых случайных величин, а в обозначении этой матрицы p означает, что при формировании в (1) величины x_i i -й $(p \times 1)$ -столбец этой матрицы –

$$\mathbf{H}_{\bullet,i}(p) = \eta_i(p) = (\eta_i(1), \eta_i(2), \dots, \eta_i(p))^T \quad (5)$$

умножается на i -ю $(p \times 1)$ -строку матрицы $\mathbf{Z}(p)$.

В (1), (2) предполагается, что в формировании текущего значения выходной переменной участвуют все p ее предыдущих значений. В общем же случае не все предыдущие значения переменной могут участвовать в этом формировании. Для того чтобы записать модель функционирования в таком общем случае, введем в рассмотрение структурные матрицы, смысл которых покажем на конкретном примере.

Пусть на текущее значение выходной переменной влияют первое, второе и четвертое ее предыдущие значения из заданного максимально возможного числа влияющих предыдущих значений $p = 5$. Тогда вместо строки

$\mathbf{Z}_{i,\bullet}^*(p)$ в (1) следует записать произведение

$$\mathbf{Z}_{i,\bullet}^*(p) \mathbf{S} = \begin{pmatrix} * & * & * & * & * \\ x_{i-1}, x_{i-2}, x_{i-3}, x_{i-4}, x_{i-5} \end{pmatrix} \times \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{pmatrix} * & * & * \\ x_{i-1}, x_{i-2}, x_{i-4} \end{pmatrix}, \quad (6)$$

где (5×3) -матрица \mathbf{S} представляет собой структурную матрицу, отражающую влияние первого, второго и четвертого предыдущих значений переменной на ее текущее значение.

Априорная информация о значении p и о том, какие именно предыдущие значения определяют текущее значение переменной в законе функционирования объекта (1) и (2), представляется структурной $(p \times m)$ -матрицей \mathbf{S} (m – число столбцов в матрице \mathbf{S} , равное числу неизвестных коэффициентов в модели). Будем предполагать, что эта структурная матрица задана.

С учетом введенной структурной матрицы закон функционирования (1) для общего случая формирования выходной переменной запишем в виде

$$x_i^* = \mathbf{Z}_{i,\bullet}^*(p) \mathbf{S} \mathbf{\theta}^0 + \mathbf{Z}_{i,\bullet}^*(p) \mathbf{S} \boldsymbol{\eta}_i(m) + \zeta_{i-1}, \quad (7)$$

где $\mathbf{\theta}^0$ – $(m \times 1)$ -вектор неизвестных детерминированных коэффициентов.

Введем обозначения

$$\bar{x}_i = \mathbf{Z}_{i,\bullet}^*(p) \mathbf{S} \mathbf{\theta}^0, \quad \varphi_{i-1} = \mathbf{Z}_{i,\bullet}^*(p) \mathbf{S} \boldsymbol{\eta}_i(m); \quad (8)$$

учитывая (8), запишем (7)

$$x_i^* = \bar{x}_i + \varphi_{i-1} + \zeta_{i-1}, \quad i = 1, 2, \dots, n. \quad (9)$$

Введем обозначения

$$\mathbf{x} = (x_{i-1}^*, x_{i-2}^*, \dots, x_{i-p}^*)^T, \quad (10)$$

$$\bar{\mathbf{x}} = (\bar{x}_{i-1}, \bar{x}_{i-2}, \dots, \bar{x}_{i-p})^T,$$

$$\boldsymbol{\zeta}(-1) = (\zeta_0, \zeta_1, \dots, \zeta_{n-1})^T, \quad (11)$$

$$\boldsymbol{\varphi}(-1) = (\varphi_0, \varphi_1, \dots, \varphi_{n-1})^T,$$

где $\boldsymbol{\zeta}(-1)$, $\boldsymbol{\varphi}(-1)$ – ненаблюдаемые случайные $(n \times 1)$ -векторы; -1 означает, что величина $\zeta_i(-1) = \zeta_{i-1}$ в (1) и (7), а величина $\varphi_i(-1) = \varphi_{i-1}$ в (9) аддитивно участвуют при формировании величины x_i^* .

Учитывая (10) и (11), запишем модель функционирования в векторном виде

$$\mathbf{x}^* = \bar{\mathbf{x}} + \boldsymbol{\varphi}(-1) + \boldsymbol{\zeta}(-1), \quad \bar{\mathbf{x}} = \mathbf{Z}(p) \mathbf{S} \mathbf{\theta}^0, \quad (12)$$

где $\bar{\mathbf{x}}$ – ненаблюдаемая составляющая $(n \times 1)$ -вектора значений переменной.

Пусть для наблюдений выходной переменной объекта выполняется

$$x_i = x_i^* + \varepsilon_i, \quad i = 1, 2, \dots, n, \quad (13)$$

где x_i – наблюдаемое значение переменной, измеренное в момент времени $t = t_i$, $i = 1, 2, \dots, n$; x_i^* – ненаблюдаемое значение, формируемое согласно (1) и (7); ε_i – случайная ненаблюдаемая ошибка измерения.

Учитывая (13), получаем модель наблюдения объекта в векторном виде

$$\mathbf{x} = \mathbf{x}^* + \boldsymbol{\varepsilon}. \quad (14)$$

Сформулируем предположения о статистических свойствах случайных величин в моделях функционирования и наблюдения.

Пусть относительно $\boldsymbol{\zeta}(-1)$ в (11) выполняются следующие предположения:

$$E\{\boldsymbol{\zeta}(-1)\} = \mathbf{0}_n, \quad E\{\boldsymbol{\zeta}(-1)\boldsymbol{\zeta}^T(-1)\} = \sigma_\zeta \cdot \mathbf{I}_n, \quad (15)$$

где $E\{\cdot\}$ – знак математического ожидания по возможным реализациям вектора $\boldsymbol{\zeta}(-1)$; $\mathbf{0}_n$ – нулевой $(n \times 1)$ -вектор; σ_ζ – дисперсия случайной величины $\zeta_i(-1)$, $i = 1, 2, \dots, n$, ограниченная величина; \mathbf{I}_n – единичная $(n \times n)$ -матрица.

Пусть относительно $\boldsymbol{\eta}_i(m)$ в (5) выполняются предположения

$$E\{\boldsymbol{\eta}_i(m)\} = \mathbf{0}_m, \quad i = 1, 2, \dots, n; \quad (16)$$

$$E\{\boldsymbol{\eta}_i(m) \boldsymbol{\eta}_i^T(m)\} = \boldsymbol{\Sigma}_\eta = \text{diag}\{\sigma_\eta(1), \sigma_\eta(2), \dots, \sigma_\eta(m)\}, \quad (17)$$

где $E\{\cdot\}$ – знак математического ожидания по возможным реализациям вектора $\boldsymbol{\eta}_i(m)$; $\mathbf{0}_m$ – нулевой $(m \times 1)$ -вектор; $\sigma_\eta(j)$ – дисперсия флуктуаций j -го коэффициента в (5), $j = 1, 2, \dots, m$; $\boldsymbol{\Sigma}_\eta$ – диагональная $(m \times m)$ -матрица.

Пусть относительно $\boldsymbol{\varepsilon}$ выполняются предположения

$$E\{\boldsymbol{\varepsilon}\} = \mathbf{0}_n, \quad E\{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T\} = \sigma_\varepsilon \cdot \mathbf{I}_n, \quad (18)$$

где $E\{\cdot\}$ – знак математического ожидания по возможным реализациям вектора $\boldsymbol{\varepsilon}$; σ_ε – дисперсия величины $\varepsilon_j, j=1, 2, \dots, m$, ограниченная величина.

Предположим, что $\sigma_\zeta, \sigma_\eta(1), \sigma_\eta(2), \dots, \sigma_\eta(m), \sigma_\varepsilon$ – дисперсии случайных величин в моделях функционирования и наблюдения – заданы.

Будем также предполагать, что случайные векторы $\zeta(-1), \boldsymbol{\varepsilon}$ и $\boldsymbol{\eta}_i(m), i=1, 2, \dots, n$, статистически независимы:

$$\begin{aligned} E\{\zeta(-1)\boldsymbol{\varepsilon}^T\} &= \mathbf{O}_{n \times n}, \quad E\{\boldsymbol{\eta}_i(m)\zeta^T(-1)\} = \mathbf{O}_{(m \times n)}, \\ E\{\boldsymbol{\eta}_i(m)\boldsymbol{\varepsilon}^T\} &= \mathbf{O}_{(m \times n)}, \end{aligned} \quad (19)$$

где $\mathbf{O}_{(n \times n)}$ – нулевая $(n \times n)$ -матрица; $\mathbf{O}_{(m \times n)}$ – нулевая $(m \times n)$ -матрица.

Пусть в результате наблюдения в моменты времени $t=t_i, i=1-2p, 2-2p, \dots, 0, 1, 2, \dots, n$ получен $(n+2p)$ -вектор значений выходной переменной

$$\left(x_{1-2p}, x_{2-2p}, \dots, x_0, x_1, x_2, \dots, x_n\right)^T = \begin{pmatrix} \mathbf{x}(0) \\ \mathbf{x} \end{pmatrix}, \quad (20)$$

где $(2p \times 1)$ -вектор $\mathbf{x}(0)$ будет использован в качестве начальных условий.

Пусть заданы:

- p – число предыдущих значений выходной переменной, влияющих на их текущее значение;
- структурная $(p \times m)$ -матрица \mathbf{S} , указывающая, какие именно предыдущие значения переменной определяют ее текущее значение объекта (1)–(19).

Для оценивания неизвестных коэффициентов $\overset{\circ}{\boldsymbol{\theta}}$ по наблюдениям объекта (20) используем результаты [17], где разработана итерационная процедура параметрической идентификации для системы авторегрессионных уравнений со случайными коэффициентами.

Оценивание параметров в авторегрессионном уравнении со случайными коэффициентами

Из модели функционирования (12) следует

$$\overset{*}{\mathbf{Z}}(p) = \overline{\mathbf{Z}}(p) + \Phi(-2, Z) + \Gamma(-2, Z), \quad (21)$$

где $\overline{\mathbf{Z}}(p)$ – $(n \times p)$ -матрица ненаблюдаемых значений переменной объекта, по своей структуре аналогичная матрице $\overset{*}{\mathbf{Z}}(p)$ в (1) и (2):

$$\overline{\mathbf{Z}}(p) = \begin{bmatrix} x_0 & x_{-1} & \cdots & x_{1-p} \\ x_1 & x_0 & \cdots & x_{2-p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{i-1} & x_{i-2} & \cdots & x_{i-p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n-1} & x_{n-2} & \cdots & x_{n-p} \end{bmatrix}; \quad (22)$$

$\Phi(-2, Z)$ – $(n \times p)$ -матрица ненаблюдаемых случайных величин

$$\Phi(-2, Z) = \begin{bmatrix} \varphi_{-1} & \varphi_{-2} & \cdots & \varphi_{-p} \\ \varphi_0 & \varphi_{-1} & \cdots & \varphi_{1-p} \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_{i-2} & \varphi_{i-3} & \cdots & \varphi_{i-1-p} \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_{n-2} & \varphi_{n-3} & \cdots & \varphi_{n-1-p} \end{bmatrix}, \quad (23)$$

в обозначении которой -2 означает, что в (21)

при формировании величины x_{i-1} аддитивно участвует величина φ_{i-2} ; $\Gamma(-2, Z)$ – $(n \times p)$ -матрица ненаблюдаемых случайных величин

$$\Gamma(-2, Z) = \begin{bmatrix} \zeta_{-1} & \zeta_{-2} & \cdots & \zeta_{-p} \\ \zeta_0 & \zeta_{-1} & \cdots & \zeta_{1-p} \\ \vdots & \vdots & \ddots & \vdots \\ \zeta_{i-2} & \zeta_{i-3} & \cdots & \zeta_{i-1-p} \\ \vdots & \vdots & \ddots & \vdots \\ \zeta_{n-2} & \zeta_{n-3} & \cdots & \zeta_{n-1-p} \end{bmatrix}, \quad (24)$$

в обозначении которой -2 означает, что в (21)

при формировании величины x_{i-1} аддитивно участвует величина ζ_{i-2} .

Подставим в (14) вектор \mathbf{x} из (12) и используем (21) для $\overset{*}{\mathbf{Z}}(p)$:

$$\mathbf{x} = \overline{\mathbf{Z}}(p) \overset{\circ}{\mathbf{S}} \boldsymbol{\theta} + \boldsymbol{\xi}, \quad (25)$$

где ξ – $(n \times 1)$ -вектор ненаблюдаемых случайных аддитивных составляющих

$$\xi = \varepsilon + \Phi(-2, Z) \overset{\circ}{S} \overset{\circ}{\theta} + \Gamma(-2, Z) \overset{\circ}{S} \overset{\circ}{\theta} + \varphi(-1) + \zeta(-1). \quad (26)$$

Используя (15) – (19) и учитывая, что все случайные векторы ε , $\varphi(-1)$, $\zeta(-1)$ и случайные матрицы $\Phi(-2, Z)$, $\Gamma(-2, Z)$ имеют нулевые математические ожидания, и все эти величины статистически независимы, для математического ожидания ξ получаем

$$E\{\xi\} = \mathbf{0}_n, \quad (27)$$

где $\mathbf{0}_n$ – нулевой $(n \times 1)$ -вектор.

Введем обозначения

$$\mathbf{y} = \mathbf{x}, \quad \mathbf{R} = \overset{=}{\mathbf{Z}}(p) \mathbf{S}, \quad (28)$$

где \mathbf{R} – $(n \times m)$ -матрица регрессоров для выходной переменной.

Учитывая (28), регрессионную модель (25) записываем в виде

$$\mathbf{y} = \mathbf{R} \overset{\circ}{\theta} + \xi = \overset{\circ}{\mathbf{y}} + \xi, \quad (29)$$

где \mathbf{y} – $(n \times 1)$ -вектор наблюдаемых зашумленных значений выходной переменной; $\overset{\circ}{\mathbf{y}}$ – $(n \times 1)$ -вектор ненаблюдаемых значений; $\overset{\circ}{\theta}$ – $(m \times 1)$ -вектор неизвестных коэффициентов.

Согласно [17], для оценки коэффициентов $\overset{\circ}{\theta}$ выполняется

$$\hat{\mathbf{d}} = \mathbf{C} \mathbf{y}, \quad (30)$$

где для $(m \times n)$ -матрицы \mathbf{C} выполняется

$$\mathbf{C} = (\mathbf{R}^T \Sigma_{\xi}^{-1} \mathbf{R})^{-1} \mathbf{R}^T \Sigma_{\xi}^{-1}; \quad (31)$$

Σ_{ξ} – ковариационная $(n \times n)$ -матрица введенного в (26) $(n \times 1)$ -вектора ненаблюдаемой аддитивной случайной составляющей ξ .

Для ковариационной матрицы выполняется соотношение из [17]:

$$\Sigma_{\xi} = \sigma_{\varepsilon} \cdot \mathbf{I}_n + \Lambda(\Phi) + \Psi + \Lambda(-1, \varphi) + \sigma_{\zeta} \cdot \mathbf{I}_n, \quad (32)$$

где σ_{ε} , σ_{ζ} – дисперсии случайных величин в

моделях функционирования и наблюдения, введенные в (18) и (15) соответственно; $\Lambda(\Phi)$ – диагональная $(n \times n)$ -матрица

$$\Lambda(\Phi) = \text{diag}[\lambda_{11}(\Phi), \lambda_{22}(\Phi), \dots, \lambda_{nn}(\Phi)], \quad (33)$$

элементы которой определены в формуле (59) в [17]; $\Lambda(-1, \varphi)$ – диагональная $(n \times n)$ -матрица

$$\Lambda(-1, \varphi) = \text{diag}[\sigma_{11}(-1, \varphi), \sigma_{22}(-1, \varphi), \dots, \sigma_{nn}(-1, \varphi)], \quad (34)$$

элементы которой вычисляются по формулам (55) и (56) в [17].

В (32) Ψ – $(n \times n)$ -матрица вида:

$$\Psi = \begin{bmatrix} \psi(0) & \psi(+1) & \dots & \psi(p-1) & 0 & \dots & 0 & 0 \\ \psi(-1) & \psi(0) & \dots & \psi(p-2) & \psi(p-1) & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \psi(1-p) & \psi(2-p) & \dots & \psi(0) & \psi(+1) & \dots & 0 & 0 \\ 0 & \psi(1-p) & \dots & \psi(-1) & \psi(0) & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & \psi(0) & \psi(+1) \\ 0 & 0 & \dots & 0 & 0 & \dots & \psi(-1) & \psi(0) \end{bmatrix}, \quad (35)$$

где величины $\psi(\Delta)$, $\Delta = -p+1, -p+2, \dots, -1, 0, 1, \dots, p-2, p-1$, определены в формулах (59) – (62) в статье [17].

С учетом (31) – (35) для оценок коэффициентов выполняется

$$\hat{\mathbf{d}} = (\mathbf{R}^T \Sigma_{\xi}^{-1} \mathbf{R})^{-1} \mathbf{R}^T \Sigma_{\xi}^{-1} \mathbf{y}. \quad (36)$$

В формулу (31) для матрицы \mathbf{C} входит ненаблюдаемая матрица регрессоров \mathbf{R} , а в формулу (32) для матрицы Σ_{ξ} – матрица Ψ , элементы которой, как следует из (33) – (49), зависят от неизвестных коэффициентов $\overset{\circ}{\theta}$. Эти обстоятельства использованы в [17] для построения итерационной процедуры вычисления неизвестных коэффициентов в виде (30) для случая, когда дисперсии σ_{ζ} , $\sigma_{\eta}(1)$, $\sigma_{\eta}(2)$, ..., $\sigma_{\eta}(m)$, σ_{ε} – дисперсии случайных величин в моделях функционирования и наблюдения – априорно известны. Процедура исследована методом статистических испытаний.

Для оценки $\hat{\mathbf{d}}$ с учетом (29) и (36) получим

$$\hat{\mathbf{d}} = [(\mathbf{R}^T \Sigma_{\xi}^{-1} \mathbf{R})^{-1}] \mathbf{R}^T \Sigma_{\xi}^{-1} \mathbf{R} \hat{\boldsymbol{\theta}} + \mathbf{C} \hat{\xi} = \hat{\boldsymbol{\theta}} + \mathbf{C} \hat{\xi}. \quad (37)$$

С учетом (37) регрессионную модель запишем в виде

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{u} = \mathbf{R} \hat{\mathbf{d}} + \mathbf{u}, \quad (38)$$

где $\hat{\mathbf{y}}$ – $(n \times 1)$ -вектор выхода регрессионной модели

$$\hat{\mathbf{y}} = \mathbf{R} \hat{\mathbf{d}} = \mathbf{R} \hat{\boldsymbol{\theta}} + \mathbf{R} \mathbf{C} \hat{\xi} = \hat{\mathbf{y}} + \mathbf{R} \mathbf{C} \hat{\xi}; \quad (39)$$

\mathbf{u} – $(n \times 1)$ -вектор остатков [18], для которого выполняется

$$\mathbf{u} = \hat{\xi} - \mathbf{R} \mathbf{C} \hat{\xi}, \quad E\{\mathbf{u}\} = \mathbf{0}_n, \quad (40)$$

т.е. его математическое ожидание равно нулевому $(n \times 1)$ -вектору.

Критерий регулярности МГУА для линейной авторегрессии со случайными коэффициентами

Пусть структурная $(p \times m)$ -матрица $\hat{\mathbf{S}}$ соответствует истинной структуре модели объекта (1) – (19), т.е. однозначно задает структуру авторегрессионного уравнения: указывает, какие именно предыдущие значения определяют текущие значения выходной переменной. Пусть эта структурная матрица неизвестна и требуется ее определить по результатам наблюдения (20), т.е. рассмотреть задачу структурной идентификации. Далее будем предполагать, что для генерации и анализа структур моделей применяется алгоритм полного перебора всех возможных структур моделей, а значение p – число предыдущих значений выходной переменной, влияющих на ее текущее значение – априорно известно. Пусть структурная $(p \times s)$ -матрица \mathbf{S} соответствует текущей анализируемой структуре модели.

Рассмотрим так называемый J -функционал качества регрессионного уравнения, отражающий требование минимизации математического ожидания

$$J = E\{(\hat{\mathbf{y}} - \mathbf{y})^T (\hat{\mathbf{y}} - \mathbf{y})\}, \quad (41)$$

известный для модели одномерной по выходу регрессии [19] и получивший в МГУА назва-

ние *идеальный внешний критерий* [11]. Функционал (41) не может применяться при решении практических задач, так как содержит ненаблюдаемый вектор $\hat{\mathbf{y}}$, но может быть использован для теоретического сравнения методов оценивания, в том числе на основе метода статистических испытаний [16].

Существует ли конструктивная альтернатива ненаблюдаемому вектору $\hat{\mathbf{y}}$ в J -функционале (41), сохраняющая для соответствующего функционала свойства J -функционала? Положительный ответ на этот вопрос для систем статических (одновременных) регрессионных уравнений [20, 21] получен в рамках так называемой схемы повторных наблюдений, которая может быть реализована в условиях активного эксперимента. В этой схеме для заданного вектора значений входных переменных объекта проводится не одно, а пара независимых наблюдений выходных переменных. Первое наблюдение из этой пары участвует в формировании выборки A , другое – выборки B , и, в результате, в схеме повторных наблюдений выполняется $\mathbf{X}(A) = \mathbf{X}(B)$.

Такую схему можно реализовать для статических регрессионных моделей, но для авторегрессионных моделей она принципиально нереализуема: значение каждой из переменных множества X в силу модели (1) формируется с участием случайной ненаблюдаемой составляющей, которую «воспроизвести» нельзя. Поскольку добиться выполнения $\mathbf{X}(A) = \mathbf{X}(B)$ невозможно, остается попытка обеспечить выполнение

$$\mathbf{X}^T(A) \mathbf{X}(A) \approx \mathbf{X}^T(B) \mathbf{X}(B) \quad (42)$$

и провести исследование критерия регулярности в этих условиях.

Будем предполагать, что объект, для которого решается задача структурной идентификации, принадлежит к классу объектов, допускающих возможность неоднократного наблюдения реализаций функционирования (например, временные ряды показателей некоторого технологического процесса). Будем выбирать такие реализации, которые начинаются с при-

близительно одинаковых начальных условий $\mathbf{x}^T(A, 0) \mathbf{x}(A, 0) \approx \mathbf{x}^T(B, 0) \mathbf{x}(B, 0)$, имеют качественно одинаковый характер переходных процессов и заканчиваются близкими состояниями в конечные моменты времени.

Пусть в качестве двух выборок наблюдений A и B выбраны наблюдения двух реализаций функционирования объекта – две выборки наблюдений выходной переменной, обладающих указанными свойствами. Первую выборку A будем называть обучающей, а вторую B – проверочной. На обучающей выборке будем оценивать коэффициенты в авторегрессионном уравнении с текущей анализируемой структурой, а на проверочной – качество построенной модели. В дальнейшем будем называть такой способ формирования обучающей и проверочной выборок «схемой квазиповторных наблюдений».

В соответствии с (40) для $(n(B) \times 1)$ -вектора остатков на выборке B выполняется

$$\begin{aligned} \mathbf{u}(B | A, S) &= \mathbf{y}(B) - \hat{\mathbf{y}}(B | A, S) = \\ &= \mathbf{y}(B) - \mathbf{R}(B, S) \hat{\mathbf{d}}(A, S), \end{aligned} \quad (43)$$

где $\mathbf{y}(B)$ – $(n(B) \times 1)$ -вектор наблюдений выходной переменной выборки B ; $\hat{\mathbf{y}}(B | A, S)$ – $(n(B) \times 1)$ -вектор выходов регрессионной модели на выборке B , рассчитанный по модели, оценки коэффициентов которой $\hat{\mathbf{d}}(A, S)$ получены в соответствии с (37) – (43) на обучающей выборке A для структуры S ; $n(B)$ – объем проверочной выборки.

В соответствии с (39) – (41) для $\mathbf{u}(B | A, S)$ выполняется

$$\begin{aligned} \mathbf{u}(B | A, S) &= \overset{\circ}{\mathbf{y}}(B) + \xi(B) - \mathbf{R}(B, S) \mathbf{C}(A, S) \overset{\circ}{\mathbf{y}}(A) - \\ &- \mathbf{R}(B, S) \mathbf{C}(A, S) \xi(A) = \\ &= \delta(B | A, S) + \xi(B) - \mathbf{R}(B, S) \mathbf{C}(A, S) \xi(A), \end{aligned} \quad (44)$$

где

$$\delta(B | A, S) = \overset{\circ}{\mathbf{y}}(B) - \mathbf{R}(B, S) \mathbf{C}(A, S) \overset{\circ}{\mathbf{y}}(A) \quad (45)$$

– $(n \times 1)$ -вектор смещения, обусловленный выбором структуры S вместо $\overset{\circ}{S}$.

Определение 1. Случайная величина

$$ARD^*(S) = \mathbf{u}^T(B | A, S) \mathbf{u}(B | A, S) \quad (46)$$

называется критерием регулярности для авторегрессионного уравнения, где звездочка означает, что выборки A и B получены в условиях квазиповторных наблюдений.

Определение 2. Оптимальным множеством регрессоров называется множество регрессоров, соответствующее набору структурных матриц S_0 :

$$S_0 = \arg \min_{S \in S^*(p)} E\{ARD^*(S)\}, \quad (47)$$

где $S^*(p)$ – множество возможных структурных матриц при заданном p .

Определение 3. Оптимальным по количеству и составу регрессоров называется авторегрессионное уравнение, построенное на множестве регрессоров, соответствующем структурной матрице S_0 .

Для математического ожидания $ARD^*(S)$ с учетом результатов [21] выполняется

$$\begin{aligned} E\{ARD^*(S)\} &= \\ &= \delta^T(B | A, S) \delta(B | A, S) + E\{\xi^T(B) \xi(B)\} - \\ &- E\{\xi^T(B) \mathbf{R}(B, S) \mathbf{C}(A, S) \xi(A)\} - \\ &- E\{\mathbf{R}(B, S) \mathbf{C}(A, S) \xi(A)\}^T \xi(B)\} + \\ &+ E\{\mathbf{R}(B, S) \mathbf{C}(A, S) \xi(A)\}^T \times \\ &\times \mathbf{R}(B, S) \mathbf{C}(A, S) \xi(A)\}. \end{aligned} \quad (48)$$

Для второго слагаемого в (48) с учетом (32) – (35) выполняется

$$\begin{aligned} E\{\xi^T(B) \xi(B)\} &= n(B) \cdot (\sigma_\varepsilon + \psi(0, S) + \sigma_\zeta) + \\ &+ \sum_{i=1}^{n(B)} \lambda_{ii}(\Phi, S) + \sum_{i=1}^{n(B)} \sigma_{ii}(-1, \varphi, S), \end{aligned} \quad (49)$$

а третье и четвертое слагаемые равны нулю в силу независимости $\xi(A)$ и $\xi(B)$.

Для пятого слагаемого в (48), учитывая (31), получаем

$$\begin{aligned} E\{\mathbf{R}(B, S) \mathbf{C}(A, S) \xi(A)\}^T \mathbf{R}(B, S) \mathbf{C}(A, S) \xi(A)\} &= \\ &= \text{tr}[\mathbf{R}(B, S) \times \\ &\times (\mathbf{R}^T(A, S) \Sigma_\xi^{-1} \mathbf{R}(A, S))^{-1} \mathbf{R}^T(B, S)]. \end{aligned} \quad (50)$$

Подставляя в (48) выражения (49) и (50), получаем

$$E\{ARD^*(S)\} = \Delta(B|A, S) + n(B) \cdot (\sigma_\varepsilon + \psi(0, S) + \sigma_\zeta) + \text{tr}[\Lambda(\Phi, \varphi, S)] + \text{tr}[\mathbf{P}(B|A, S)], \quad (51)$$

где $\Delta(B|A, S)$ – смещение, обусловленное выбором структуры S вместо $\overset{\circ}{S}$:

$$\Delta(B|A, S) = \delta^T(B|A, S)\delta(B|A, S); \quad (52)$$

$$\Lambda(\Phi, \varphi, S) – \text{диагональная } (n(B) \times n(B))\text{-матрица}$$

$$[\Lambda(\Phi, \varphi, S)] = \text{diag}[\lambda_{11}(\Phi, S) + \sigma_{11}(-1, \varphi, S), \dots, \lambda_{mm}(\Phi, S) + \sigma_{mm}(-1, \varphi, S)]; \quad (53)$$

$$\mathbf{P}(B|A, S) = \mathbf{R}(B, S)(\mathbf{R}^T(A, S)\Sigma_\xi^{-1}\mathbf{R}(A, S))^{-1}\mathbf{R}^T(B, S). \quad (54)$$

Исследование критерия регулярности МГУА

Установим свойства критерия регулярности МГУА (46). С этой целью исследуем, как изменяется математическое ожидание критерия в зависимости от состава множества регрессоров.

В случае истинной структуры $\overset{\circ}{S}$ для математического ожидания критерия регулярности выполняется

$$E\{ARD^*(\overset{\circ}{S})\} = n(B) \cdot (\sigma_\varepsilon + \psi(0, \overset{\circ}{S}) + \sigma_\zeta) + \text{tr}[\Lambda(\Phi, \varphi, \overset{\circ}{S})] + \text{tr}[\mathbf{P}(B|A, \overset{\circ}{S})], \quad (55)$$

где матрицы $\Lambda(\Phi, \varphi, \overset{\circ}{S})$, $\mathbf{P}(B|A, \overset{\circ}{S})$ могут быть записаны аналогично (53), (54).

Случай недостающего регрессора. Рассмотрим случай, когда в текущей структуре пропущен один регрессор. Предположим для простоты, что это регрессор с номером p , он – максимально удален предыдущим значением переменной, участвующим в формировании ее текущего значения, т.е. для структурных матриц выполняется

$$\overset{\circ}{S} = [\mathbf{S} \mid \mathbf{s}], \quad (56)$$

где $\overset{\circ}{S}$ – структурная $(p \times m)$ -матрица истинной модели; \mathbf{S} – структурная $(p \times (m-1))$ -матрица текущей модели; для \mathbf{s} – $(p \times 1)$ -вектора выполняется

$$\mathbf{s} = (0, 0, \dots, 0, 1)^T. \quad (57)$$

Другими словами, в модели функционирования (4) в формировании величины x_i участвует величина x_{i-p} , но в текущую модель она не включена.

Для матриц регрессоров, соответствующих $\overset{\circ}{S}$ и S в (56), выполняется

$$\mathbf{R}(\overset{\circ}{S}) = \mathbf{X}\overset{\circ}{S} = \mathbf{X}[\mathbf{S} \mid \mathbf{s}] = [\mathbf{X}\mathbf{S} \mid \mathbf{X}\mathbf{s}] = [\mathbf{R}(S) \mid \mathbf{m}], \quad (58)$$

где \mathbf{m} – $(n \times 1)$ -вектор наблюдений пропущенного регрессора.

Используя (51) – (55), вычислим разность математических ожиданий критерия регулярности для текущей структуры S и истинной структуры $\overset{\circ}{S}$:

$$\Delta_1(S, \overset{\circ}{S}) = E\{ARD^*(S)\} - E\{ARD^*(\overset{\circ}{S})\} = \Delta(B|A, S) + n(B) \cdot \psi(0, S) - n(B) \cdot \psi(0, \overset{\circ}{S}) + \text{tr}[\Lambda(\Phi, \varphi, S)] - \text{tr}[\Lambda(\Phi, \varphi, \overset{\circ}{S})] + \text{tr}[\mathbf{P}(B|A, S)] - \text{tr}[\mathbf{P}(B|A, \overset{\circ}{S})], \quad (59)$$

где $\Delta(B|A, S) = \delta^T(B|A, S)\delta(B|A, S)$; $\delta(B|A, S)$ – введенный в (45) $(n \times 1)$ -вектор смещения, обусловленный выбором структуры S вместо $\overset{\circ}{S}$.

Вычисление разности (59) состоит из ряда шагов.

Шаг 1. Для вектора смещения $\delta(B|A, S)$ выполняется

$$\delta(B|A, S) = \mathbf{R}(B, \overset{\circ}{S})\overset{\circ}{\theta} - \mathbf{R}(B, S)(\mathbf{R}^T(A, S)\Sigma_\xi^{-1}\mathbf{R}(A, S))^{-1} \times \mathbf{R}^T(A, S)\Sigma_\xi^{-1}\mathbf{R}(A, \overset{\circ}{S})\overset{\circ}{\theta}. \quad (60)$$

Запишем (60) с учетом (58)

$$\delta(B|A, S) = [\mathbf{R}(B, S) \mid \mathbf{m}(B)]\overset{\circ}{\theta} - [\mathbf{R}(B, S) \mid \mathbf{R}(B, S)(\mathbf{R}^T(A, S)\Sigma_\xi^{-1} \times$$

$$\times \mathbf{R}(A, S))^{-1} \mathbf{R}^T(A, S) \Sigma_{\xi}^{-1} \mathbf{m}(A)] \overset{\circ}{\boldsymbol{\theta}} =$$

$$= [\mathbf{O}_{n \times (m-1)} \quad | \quad \mathbf{m}(B) - \mathbf{P}(B | A, S) \Sigma_{\xi}^{-1} \mathbf{m}(A)] \overset{\circ}{\boldsymbol{\theta}}, \quad (61)$$

$$\mathbf{P}(B | A, S) = \mathbf{R}(B, S) \times$$

$$\times (\mathbf{R}^T(A, S) \Sigma_{\xi}^{-1} \mathbf{R}(A, S))^{-1} \mathbf{R}^T(A, S). \quad (62)$$

Учитывая (61) и (62) для величины $\Delta(B | A, S)$ в (59), получаем

$$\Delta(B | A, S) =$$

$$= \overset{\circ}{\boldsymbol{\theta}}^T \left[\begin{array}{c|c} \mathbf{O}_{(m-1) \times (m-1)} & \mathbf{0}_{m-1} \\ \hline \mathbf{0}_{m-1}^T & (\mathbf{m}^T(B) - \mathbf{m}^T(A) \Sigma_{\xi}^{-1} \mathbf{P}^T(B | A, S)) \times \\ & \times (\mathbf{m}(B) - \mathbf{P}(B | A, S) \Sigma_{\xi}^{-1} \mathbf{m}(A)) \end{array} \right] \overset{\circ}{\boldsymbol{\theta}}. \quad (63)$$

Для (m, m) -го элемента матрицы в (63) выполняется

$$a = \mathbf{m}^T(B) \mathbf{m}(B) - \mathbf{m}^T(B) \mathbf{R}(B, S) \times$$

$$\times (\mathbf{R}^T(A, S) \Sigma_{\xi}^{-1} \mathbf{R}(A, S))^{-1} \mathbf{R}^T(A, S) \Sigma_{\xi}^{-1} \mathbf{m}(A) -$$

$$- \mathbf{m}^T(A) \Sigma_{\xi}^{-1} \mathbf{R}(A, S) (\mathbf{R}^T(A, S) \times$$

$$\times \Sigma_{\xi}^{-1} \mathbf{R}(A, S))^{-1} \mathbf{R}^T(B, S) \mathbf{m}(B) +$$

$$+ \mathbf{m}^T(A) \Sigma_{\xi}^{-1} \mathbf{R}(A, S) (\mathbf{R}^T(A, S) \times$$

$$\times \Sigma_{\xi}^{-1} \mathbf{R}(A, S))^{-1} \mathbf{R}^T(B, S) \times$$

$$\times \mathbf{R}(B, S) (\mathbf{R}^T(A, S) \Sigma_{\xi}^{-1} \mathbf{R}(A, S))^{-1} \times$$

$$\times \mathbf{R}^T(A, S) \Sigma_{\xi}^{-1} \mathbf{m}(A). \quad (64)$$

Далее используем то обстоятельство, что обучающая (A) и проверочная (B) выборки получены особым способом как пара реализаций функционирования объекта с близкими начальными условиями, качественно одинаковым характером переходных процессов и близкими состояниями в конечные моменты времени. Введем в рассмотрение разности:

$$\mathbf{G}(X) = \mathbf{X}^T(B) \mathbf{X}(B) - \mathbf{X}^T(A) \mathbf{X}(A), \quad (65)$$

$$g(\mathbf{m}) = \mathbf{m}^T(B) \mathbf{m}(B) - \mathbf{m}^T(A) \mathbf{m}(A), \quad (66)$$

$$\mathbf{G}(R, S) = \mathbf{R}^T(B, S) \mathbf{R}(B, S) - \mathbf{R}^T(A, S) \mathbf{R}(A, S), \quad (67)$$

$$g^T(R, S) = \mathbf{m}^T(B) \mathbf{R}(B, S) - \mathbf{m}^T(A) \mathbf{R}(A, S). \quad (68)$$

Оценивание влияния скаляра $g(\mathbf{m})$, вектора $g(R, S)$, матрицы $\mathbf{G}(R, S)$ на условие ре-

дукции (упрощения) структуры по числу регрессоров проведено методом статистических испытаний в отдельном исследовании [22].

Учитывая (65) – (68), для (64) получаем

$$a = \mathbf{m}^T(A) [\mathbf{M}^T(A, S) \mathbf{M}(A, S)] \mathbf{m}(A) + \rho_1(\mathbf{m}), \quad (69)$$

где $\mathbf{M}(A, S)$ – $(n(A) \times n(A))$ -матрица

$$\mathbf{M}(A, S) = [\mathbf{I}_{n(A)} - \mathbf{R}(A, S) (\mathbf{R}^T(A, S) \times$$

$$\times \Sigma_{\xi}^{-1} \mathbf{R}(A, S))^{-1} \mathbf{R}^T(A, S) \Sigma_{\xi}^{-1}], \quad (70)$$

а для скаляра $\rho_1(\mathbf{m})$ выполняется

$$\rho_1(\mathbf{m}) = g(\mathbf{m}) - g^T(R, S) \mathbf{C}(A, S) \mathbf{m}(A) -$$

$$- \mathbf{m}^T(A) \mathbf{C}^T(A, S) g(R, S) +$$

$$+ \mathbf{m}^T(A) \mathbf{C}^T(A, S) \mathbf{G}(R, S) \mathbf{C}(A, S) \mathbf{m}(A). \quad (71)$$

Учитывая (56), (69), (70) и соотношение

$$\overset{\circ}{\boldsymbol{\theta}}^T = (\overset{\circ}{\theta}_1, \overset{\circ}{\theta}_2, \dots, \overset{\circ}{\theta}_m), \quad (72)$$

получаем

$$\Delta(B | A, S) =$$

$$= \overset{\circ}{\boldsymbol{\theta}}^T \left[\begin{array}{c|c} \mathbf{O}_{(m-1) \times (m-1)} & \mathbf{0}_{m-1} \\ \hline \mathbf{0}_{m-1}^T & \mathbf{m}^T(A) \mathbf{M}^T(A, S) \mathbf{M}(A, S) \mathbf{m}(A) + \rho_1(\mathbf{m}) \end{array} \right] \overset{\circ}{\boldsymbol{\theta}} =$$

$$= (\overset{\circ}{\theta}_m)^2 \cdot \mathbf{m}^T(A) \mathbf{H}(A, S) \mathbf{m}(A) + (\overset{\circ}{\theta}_m)^2 \cdot \rho_1(\mathbf{m}), \quad (73)$$

$$\mathbf{H}(A, S) = \mathbf{M}^T(A, S) \mathbf{M}(A, S). \quad (74)$$

Итак, в (60) – (74) установлено

$$\Delta(B | A, S) = \delta^T(B | A, S) \delta(B | A, S) =$$

$$= (\overset{\circ}{\theta}_m)^2 \cdot \mathbf{m}^T(A) \mathbf{H}(A, S) \mathbf{m}(A) + (\overset{\circ}{\theta}_m)^2 \cdot \rho_1(\mathbf{m}). \quad (75)$$

Шаг 2. Для скалярных величин $n(B) \cdot \psi(0, S)$ и $n(B) \cdot \psi(0, \overset{\circ}{S})$ в (59) с учетом формул (59), (60) в статье [17] выполняется

$$\Delta(\psi) = n(B) \cdot \psi(0, S) - n(B) \cdot \psi(0, \overset{\circ}{S}) =$$

$$= -n(B) \cdot \sigma_{\xi} \cdot (\overset{\circ}{\theta}_m)^2. \quad (76)$$

Шаг 3. Для следов $\text{tr}[\Lambda(\Phi, \varphi, S)]$ и $\text{tr}[\Lambda(\Phi, \varphi, \overset{\circ}{S})]$ в (59) с учетом формул (55), (56) и (59) в [17] выполняется

$$\text{tr}[\Lambda(\Phi, \varphi, S)] - \text{tr}[\Lambda(\Phi, \varphi, \overset{\circ}{S})] =$$

$$= -\sum_{i=1}^n (\overset{\circ}{\theta}_m)^2 \cdot \overline{\mathbf{Z}}_{i-1-p, p}(-1) \cdot \overline{\mathbf{Z}}_{i-1-p, p}(-1). \quad (77)$$

Шаг 4. Для разности $\text{tr}[\mathbf{P}(B|A,S)] - \text{tr}[\mathbf{P}(B|A,\overset{\circ}{S})]$ в (59), учитывая (58) и применяя формулу обращения блочной матрицы (частный случай формулы Фробениуса [19]) для перемножения матриц, получаем:

$$\begin{aligned} & \mathbf{P}(B|A,S) - \mathbf{P}(B|A,\overset{\circ}{S}) = \\ & = \mathbf{R}(B,S)[\mathbf{R}^T(A,S)\Sigma_{\xi}^{-1}\mathbf{R}(A,S)]^{-1}\mathbf{R}^T(B,S) - \\ & \quad - [\mathbf{R}(B,S) | \mathbf{m}(B)] \times \\ & \quad \times \left[\begin{array}{c|c} \mathbf{R}^T(A,S)\Sigma_{\xi}^{-1}\mathbf{R}(A,S) & \mathbf{R}^T(A,S)\Sigma_{\xi}^{-1}\mathbf{m}(A) \\ \hline \mathbf{m}^T(A)\Sigma_{\xi}^{-1}\mathbf{R}(A,S) & \mathbf{m}^T(A)\Sigma_{\xi}^{-1}\mathbf{m}(A) \end{array} \right]^{-1} \times \\ & \quad \times \left[\begin{array}{c} \mathbf{R}^T(B,S) \\ \hline \mathbf{m}^T(B) \end{array} \right] = \\ & = -f^{-1} \cdot \mathbf{R}(B,S)\mathbf{B}^{-1}\mathbf{R}^T(A,S)\Sigma_{\xi}^{-1}\mathbf{m}(A) \times \\ & \quad \times \mathbf{m}^T(A)\Sigma_{\xi}^{-1}\mathbf{R}(A,S)\mathbf{B}^{-1}\mathbf{R}^T(B,S) + \\ & + f^{-1} \cdot \mathbf{R}(B,S)\mathbf{B}^{-1}\mathbf{R}^T(A,S)\Sigma_{\xi}^{-1}\mathbf{m}(A)\mathbf{m}^T(B) + \\ & \quad + f^{-1} \cdot \mathbf{m}(B)\mathbf{m}^T(A)\Sigma_{\xi}^{-1}\mathbf{R}(A,S) \times \\ & \quad \times \mathbf{B}^{-1}\mathbf{R}^T(B,S) - f^{-1} \cdot \mathbf{m}(B)\mathbf{m}^T(B), \quad (78) \end{aligned}$$

где

$$\begin{aligned} f & = \mathbf{m}^T(A)\Sigma_{\xi}^{-1}\mathbf{m}(A) - \\ & - \mathbf{m}^T(A)\Sigma_{\xi}^{-1}\mathbf{R}(A,S)(\mathbf{R}^T(A,S) \times \\ & \times \Sigma_{\xi}^{-1}\mathbf{R}(A,S))^{-1}\mathbf{R}^T(A,S)\Sigma_{\xi}^{-1}\mathbf{m}(A) = \\ & = \mathbf{m}^T(A)\Sigma_{\xi}^{-1}\mathbf{M}(A,S)\mathbf{m}(A). \quad (79) \end{aligned}$$

Учитывая (65) – (68), для (78) получаем

$$\begin{aligned} & \text{tr}[\mathbf{P}(B|A,S) - \mathbf{P}(B|A,\overset{\circ}{S})] = \\ & = -f^{-1} \cdot \mathbf{m}^T(A)\mathbf{H}(A,S)\mathbf{m}(A) - f^{-1} \cdot q(\mathbf{m}), \quad (80) \end{aligned}$$

где $\mathbf{H}(A,S) = \mathbf{M}^T(A,S)\mathbf{M}(A,S)$; $\mathbf{M}(A,S)$ – $(n(A) \times n(A))$ -матрица введена в (69), (70), а для скалярной величины $q(\mathbf{m})$ выполняется

$$\begin{aligned} q(\mathbf{m}) & = g(\mathbf{m}) - \mathbf{g}^T(R)\mathbf{C}(A,S)\mathbf{m}(A) - \\ & - \mathbf{m}^T(A)\mathbf{C}^T(A,S)\mathbf{g}^T(R) + \\ & + \mathbf{m}^T(A)\mathbf{C}^T(A,S)\mathbf{G}(R)\mathbf{C}(A,S)\mathbf{m}(A). \quad (81) \end{aligned}$$

Объединив результаты четырех шагов, получим

$$\Delta_1(S,\overset{\circ}{S}) = (\overset{\circ}{\theta}_m)^2 \cdot \mathbf{m}^T(A)\mathbf{H}(A,S)\mathbf{m}(A) + (\overset{\circ}{\theta}_m)^2 \cdot \rho_1(\mathbf{m}) -$$

$$\begin{aligned} & - n(B) \cdot \sigma_{\zeta} \cdot (\overset{\circ}{\theta}_m)^2 - \\ & - \sum_{i=1}^n (\overset{\circ}{\theta}_m)^2 \cdot \bar{\mathbf{Z}}_{i-1-p,p}(-1) \cdot \bar{\mathbf{Z}}_{i-1-p,p}(-1) - \\ & - f^{-1} \cdot \mathbf{m}^T(A)\mathbf{H}(A,S)\mathbf{m}(A) - f^{-1} \cdot q(\mathbf{m}). \quad (82) \end{aligned}$$

Если $\Delta_1(S,\overset{\circ}{S}) > 0$, то структура $\overset{\circ}{S}$ лучше S ; если $\Delta_1(S,\overset{\circ}{S}) < 0$, то структура S лучше $\overset{\circ}{S}$; если $\Delta_1(S,\overset{\circ}{S}) = 0$, то структура S лучше $\overset{\circ}{S}$ по дополнительному принципу простоты. Выполнение $\Delta_1(S,\overset{\circ}{S}) \leq 0$ служит условием так называемой редукции (упрощения), оптимальной по структуре модели. Из (82) для условия редукции получим

$$\begin{aligned} & (\overset{\circ}{\theta}_m)^2 \cdot \mathbf{m}^T(A)\mathbf{H}(A,S)\mathbf{m}(A) + (\overset{\circ}{\theta}_m)^2 \cdot \rho_1(\mathbf{m}) < \\ & < n(B) \cdot \sigma_{\zeta} \cdot (\overset{\circ}{\theta}_m)^2 + \\ & + \sum_{i=1}^n (\overset{\circ}{\theta}_m)^2 \cdot \bar{\mathbf{Z}}_{i-1-p,p}(-1) \cdot \bar{\mathbf{Z}}_{i-1-p,p}(-1) + \\ & + f^{-1} \cdot \mathbf{m}^T(A)\mathbf{H}(A,S)\mathbf{m}(A) - f^{-1} \cdot q(\mathbf{m}), \quad (83) \end{aligned}$$

где $\mathbf{H}(A,S)$ – $(n \times n)$ -матрица введена в (80); $\rho_1(\mathbf{m})$, f и $q(\mathbf{m})$ – скалярные величины, определенные в (71), (79) и (81) соответственно.

Исследование степени влияния скалярных величин $\rho_1(\mathbf{m})$ и $q(\mathbf{m})$ на условие редукции (83) в условиях схемы квазиповторных наблюдений, проведенное методом статистических испытаний, показало, что этими величинами можно пренебречь [22]. Тогда условие редукции примет вид

$$\begin{aligned} & (\overset{\circ}{\theta}_m)^2 \cdot \mathbf{m}^T(A)\mathbf{H}(A,S)\mathbf{m}(A) < n(B) \cdot \sigma_{\zeta} \cdot (\overset{\circ}{\theta}_m)^2 + \\ & + \sum_{i=1}^n (\overset{\circ}{\theta}_m)^2 \cdot \bar{\mathbf{Z}}_{i-1-p,p}(-1) \cdot \bar{\mathbf{Z}}_{i-1-p,p}(-1) + \\ & + f^{-1} \cdot \mathbf{m}^T(A)\mathbf{H}(A,S)\mathbf{m}(A). \quad (84) \end{aligned}$$

Учитывая (79), из (84) получаем

$$\begin{aligned} & (\overset{\circ}{\theta}_m)^2 \cdot \mathbf{m}^T(A)\Sigma_{\xi}^{-1}\mathbf{M}(A,S)\mathbf{m}(A) < 1 + \\ & + \frac{\mathbf{m}^T(A)\Sigma_{\xi}^{-1}\mathbf{M}(A,S)\mathbf{m}(A)}{\mathbf{m}^T(A)\mathbf{H}(A,S)\mathbf{m}(A)} \times \end{aligned}$$

$$\begin{aligned} & \times \{ n(B) \cdot \sigma_{\zeta} \cdot (\overset{\circ}{\theta}_m^{\circ})^2 + \\ & + \sum_{i=1}^n (\overset{\circ}{\theta}_m^{\circ})^2 \cdot \bar{\bar{\mathbf{Z}}}_{i-1-p,p}(-1) \cdot \bar{\bar{\mathbf{Z}}}_{i-1-p,p}(-1) \}. \end{aligned} \quad (85)$$

Из (85) следует, что возможность редукции модели может быть обусловлена четырьмя причинами:

- малой нормой коэффициента $\overset{\circ}{\theta}_m^{\circ}$;
- малой нормой вектора наблюдений регрессора $\mathbf{m}(A)$;
- малым объемом выборок наблюдений $n(B)$;
- малым значением величины σ_{ζ} .

Редукция модели, оптимальной по составу регрессоров, означает, что при выполнении соотношения между параметрами модели (85) следует исключить из модели регрессор \mathbf{m} . Редуцированная (упрощенная) модель будет иметь меньшую ошибку прогнозирования выходных переменных на новых выборках наблюдений в сравнении с моделью с истинной структурой.

Случай избыточного регрессора. Рассмотрим случай, когда в текущую структуру включен излишний регрессор. Предположим для простоты, что это – регрессор с номером p , максимально удаленный предыдущим значением переменной и не участвует в формировании ее текущего значения, т.е. для структурных матриц выполняется

$$\mathbf{S} = [\overset{\circ}{\mathbf{S}} \mid \mathbf{s}], \quad (86)$$

где $\overset{\circ}{\mathbf{S}}$ – структурная $(p \times m)$ -матрица истинной модели; \mathbf{S} – структурная $(p \times (m+1))$ -матрица текущей модели; \mathbf{s} – $(p \times 1)$ -вектор

$$\mathbf{s} = (0, 0, \dots, 0, 1)^T. \quad (87)$$

Другими словами, в модели функционирования (7) в формировании x_i не участвует величина x_{i-p} , но в текущую модель она включена.

Сначала вычислим математическое ожидание критерия регулярности. В рассматриваемом

случае избыточного регрессора величина $\Delta(B|A, S)$ в (52) – нулевая. Для доказательства покажем, что вектор смещения $\delta(B|A, S)$ в случае избыточного регрессора – нулевой. Действительно, различие в структурных матрицах (86), (87) приводит к выполнению

$$\begin{aligned} \mathbf{R}(S) &= \mathbf{X}\mathbf{S} = \mathbf{X}[\overset{\circ}{\mathbf{S}} \mid \mathbf{s}] \\ &= [\mathbf{X}\overset{\circ}{\mathbf{S}} \mid \mathbf{X}\mathbf{s}] = [\overset{\circ}{\mathbf{R}}(S) \mid \mathbf{r}], \end{aligned} \quad (88)$$

где \mathbf{r} – $(n \times 1)$ -вектор наблюдений избыточного регрессора.

Используя (51)–(55), вычислим разность математических ожиданий критерия регулярности для текущей структуры S и истинной структуры $\overset{\circ}{S}$:

$$\begin{aligned} \Delta_2(S, \overset{\circ}{S}) &= E\{ARD^*(S)\} - E\{ARD^*(\overset{\circ}{S})\} = \\ &= \Delta(B|A, S) + \text{tr}[\mathbf{P}(B|A, S)] - \text{tr}[\mathbf{P}(B|A, \overset{\circ}{S})], \end{aligned} \quad (89)$$

где $\Delta(B|A, S) = \delta^T(B|A, S)\delta(B|A, S)$; $\delta(B|A, S)$ – введенный в (45) $(n \times 1)$ -вектор смещения, обусловленный выбором структуры S вместо $\overset{\circ}{S}$.

Для смещения $\delta(B|A, S)$ в случае избыточного регрессора выполняется

$$\begin{aligned} \delta(B|A, S) &= \mathbf{R}(B, \overset{\circ}{S})\overset{\circ}{\theta} - [\mathbf{R}(B, \overset{\circ}{S}) \mid \mathbf{r}(B)] \times \\ & \times \left(\left[\begin{array}{c} \mathbf{R}^T(A, \overset{\circ}{S}) \\ \mathbf{r}^T(A) \end{array} \right] \Sigma_{\zeta}^{-1} [\mathbf{R}(A, \overset{\circ}{S}) \mid \mathbf{r}(A)] \right)^{-1} \times \\ & \times \left[\begin{array}{c} \mathbf{R}^T(A, \overset{\circ}{S}) \\ \mathbf{r}^T(A) \end{array} \right] \Sigma_{\zeta}^{-1} \mathbf{R}(A, \overset{\circ}{S})\overset{\circ}{\theta}. \end{aligned} \quad (90)$$

Далее, проведя вычисления аналогично (60)–(75), получим $\delta(B|A, S) = \mathbf{0}$ и, соответственно, $\Delta(B|A, S) = 0$.

Для случая избыточного регрессора аналогично (78)–(81) получим

$$\begin{aligned} & \text{tr}[\mathbf{P}(B, A, S) - \mathbf{P}(B, A, \overset{\circ}{S})] = \\ & = f^{-1}(\overset{\circ}{S}) \cdot \mathbf{r}^T(A) \mathbf{M}^T(A, \overset{\circ}{S}) \mathbf{M}(A, \overset{\circ}{S}) \mathbf{r}(A) + \\ & + f^{-1}(\overset{\circ}{S}) \cdot q(\mathbf{r}), \end{aligned} \quad (91)$$

где

$$\begin{aligned} f(\overset{\circ}{S}) &= \mathbf{r}^T(A) \Sigma_{\xi}^{-1} \mathbf{r}(A) - \mathbf{r}^T(A) \Sigma_{\xi}^{-1} \mathbf{R}(\overset{\circ}{S}) \times \\ &\times (\mathbf{R}^T(\overset{\circ}{S}) \Sigma_{\xi}^{-1} \mathbf{R}(\overset{\circ}{S}))^{-1} \mathbf{R}^T(\overset{\circ}{S}) \Sigma_{\xi}^{-1} \mathbf{r}(A) = \\ &= \mathbf{r}^T(A) \Sigma_{\xi}^{-1} \mathbf{M}(A, \overset{\circ}{S}) \mathbf{r}(A) \end{aligned} \quad (92)$$

– положительная величина, поскольку матрицы Σ_{ξ}^{-1} и $\mathbf{M}(A, \overset{\circ}{S})$ положительно определены; $\mathbf{M}(A, \overset{\circ}{S})$ – $(n \times n)$ -матрица аналогична (69), (70); $\mathbf{r}(A)$ – $(n \times 1)$ -вектор, определенный в (88); для скалярной величины $q(\mathbf{r})$ выполняется

$$\begin{aligned} q(\mathbf{r}) &= g(\mathbf{r}) - \mathbf{g}^T(R) \mathbf{B}^{-1} \mathbf{R}^T(A, \overset{\circ}{S}) \Sigma_{\xi}^{-1} \mathbf{r}(A) - \\ &- \mathbf{r}^T(A) \Sigma_{\xi}^{-1} \mathbf{R}(A, \overset{\circ}{S}) \mathbf{B}^{-1} \mathbf{g}(R) + \\ &+ \mathbf{r}^T(A) \Sigma_{\xi}^{-1} \mathbf{R}(A, \overset{\circ}{S}) \mathbf{B}^{-1} \mathbf{G}(R) \mathbf{B}^{-1} \mathbf{R}^T(A, \overset{\circ}{S}) \Sigma_{\xi}^{-1} \mathbf{r}(A). \end{aligned} \quad (93)$$

Учитывая (91), пренебрегая влиянием $f^{-1}(\overset{\circ}{S}) \cdot q(\mathbf{r})$, получаем

$$\begin{aligned} \Delta_2(S, \overset{\circ}{S}) &\approx f^{-1}(\overset{\circ}{S}) \cdot \mathbf{r}^T(A) \mathbf{M}^T \times \\ &\times (A, \overset{\circ}{S}) \mathbf{M}(A, \overset{\circ}{S}) \mathbf{r}(A) > 0, \end{aligned} \quad (94)$$

где матрица $\mathbf{H}(A, \overset{\circ}{S}) = \mathbf{M}^T(A, \overset{\circ}{S}) \mathbf{M}(A, \overset{\circ}{S})$ – положительно определена, а величина $f^{-1}(\overset{\circ}{S})$ положительна, откуда следует выполнение (94). Из (94) следует, что в случае избыточного регрессора истинная структура $\overset{\circ}{S}$ всегда лучше структуры S , а регрессор \mathbf{r} действительно не следует включать в модель.

Заключение. Критерий регулярности для моделирования в классе авторегрессионных уравнений со случайными коэффициентами в условиях квазиповторных наблюдений построен и исследован по принципам МГУА. Получены условия существования оптимального множества регрессоров, зависящие от коэффициентов моделей, матриц наблюдений регрессоров, дисперсии случайных составляющих и объемов выборок. Выявлены закономерности упрощения (редукции) оптимальной регрессионной модели. При выполнении условия ре-

дукции модель будет иметь меньшую ошибку прогнозирования выходной переменной на новых выборках наблюдений в сравнении с моделью, построенной на истинной структуре.

1. *Современные методы идентификации систем.* – М.: Мир, 1983. – 400 с.
2. *Сильвестров А.Н., Чинаев П.И.* Идентификация и оптимизация автоматических систем. – М.: Энергоатомиздат, 1987. – 199 с.
3. *Льюнг Л.* Идентификация систем. Теория для пользователя. – М.: Наука, 1991. – 432 с.
4. *Green W.H.* *Econometric Analysis.* – New Jersey: Pearson Education, Inc., 2002. – 1056 p.
5. *Söderström T., Soverini U., Mahata K.* Perspectives on errors-in-variables estimation for dynamic systems // *Signal Processing.* – 2002. – **82**, N 8. – P. 1139–1154.
6. *Кунцевич В.М.* Управление в условиях неопределенности: гарантированные результаты в задачах управления и идентификации. – К.: Наук. думка, 2006. – 264 с.
7. *Markovsky I., Van Huffel S.* Overview of total least squares methods // *Signal Processing.* – 2007. – **87**. – P. 2283–2302.
8. *Söderström T.* Errors-in-variables methods in system identification // *Automatica.* – 2007. – **43** (6). – P. 939–958.
9. *Ивахненко А.Г.* Индуктивный метод самоорганизации моделей сложных систем. – Киев: Наук. думка, 1982. – 296 с.
10. *Self-organizing methods in modelling: GMDH type algorithms / Ed. by S.J. Farlow.* – New York, Basel: Marcel Decker Inc., 1984. – 350 p.
11. *Ивахненко А.Г., Степашко В.С.* Помехоустойчивость моделирования. – Киев: Наук. думка, 1985. – 216 с.
12. *Ивахненко А.Г., Мюллер Й.А.* Самоорганизация прогнозирующих моделей. – Киев: Техніка, 1985. – 223 с.
13. *Ивахненко А.Г., Юрачковский Ю.П.* Моделирование сложных систем по экспериментальным данным. – М.: Радио и связь, 1987. – 120 с.
14. *Madala H.R., Ivakhnenko A.G.* *Inductive Learning Algorithms for Complex System Modeling.* – London, Tokyo: CRC Press Inc., 1994. – 370 p.
15. *Muller J.-A., Lemke F.* *Self-organizing Data Mining. Extracting Knowledge from Data.* – Hamburg: Libri, 2000. – 250 p.
16. *Сарычев А.П.* Идентификация состояний структурнонеопределенных систем. – Днепропетровск: Ин-т техн. механики НАН и НКА Украины, 2008. – 268 с.
17. *Сарычев А.П.* Идентификация параметров систем авторегрессионных уравнений со случайными коэффициентами при известных ковариационных матрицах // *Проблемы управления и информатики.* – 2013. – № 5. – С. 33–52.
18. *Себер Дж.* Линейный регрессионный анализ. – М.: Мир, 1980. – 456 с.

19. *Ермаков С.М., Жиглявский А.А.* Математическая теория оптимального эксперимента. – М.: Наука, 1987. – 320 с.
20. *Сарычев А.П.* Моделирование в классе систем регрессионных уравнений на основе метода группового учета аргументов // Проблемы управления и информатики. – 2013. – № 2. – С. 8–24.
21. *Сарычев А.П.* Моделирование в классе систем регрессионных уравнений со случайными коэффициентами на основе метода группового учета аргу-

ментов // Индуктивное моделирование складных систем. – 2014. – 6. – С. 137–156.

22. *Сарычев А.П.* Моделирование в классе систем авторегрессионных уравнений в условиях структурной неопределенности // Проблемы управления и информатики. – 2015. – № 4. – С. 79–103.

Поступила 09.04.2015
E-mail: Sarychev@prognoz.dp.ua
© А.П. Сарычев, 2016

UDC 519.25:681.5

A.P. Sarychev

Linear Autoregression with Random Coefficients Based on the Group Method of Data Handling in Conditions of Quasirepeated Observations

Keywords: structural uncertainty, criterion of regularity.

Introduction and purpose: The linear autoregression equation is traditional mathematical object in the theory and practice of the Group Method of Data Handling (GMDH). In 80-th years of the last century academician O.G. Ivakhnenko often posed such tasks in connection with so-called “the objective system analysis (OSA)” and then, as a rule, as criterion of selection of models (parameter of quality of regression equation) the criterion of regularity of GMDH was applied. The developed criterion is the criterion of regularity which is constructed with dividing of observations on training and testing subsamples in conditions of quasirepeated observations.

Methods: Object of research is process of modelling in a class of autoregression equations in conditions of uncertainty on structure of regressors. In this theoretical article we used the multivariate statistical analysis, the regression analysis, the theory of matrixes, the mathematical analysis and the Group Method of Data Handling.

Results: For modeling in a class of autoregression equations the criterion of regularity with dividing of observation sample on training and testing subsamples in conditions of quasirepeated observations is offered. It is proved, that the optimum set of regressors exists. The condition of a reduction of optimal autoregression equation is obtained. This condition depends on parameters of autoregression equation and volumes of samples.

Conclusion: The developed criterion of regularity allows solving a problem of structural identification in a class of autoregression equations in conditions of quasirepeated observations and can be recommended at the decision of various scientific and practical problems.



Внимание !

**Оформление подписки для желающих
опубликовать статьи в нашем журнале обязательно.**

В розничную продажу журнал не поступает.

Подписной индекс 71008