УДК 519.25:681.5

А.П. Сарычев

Линейная регрессия со случайными коэффициентами на основе метода группового учета аргументов

Исследован критерий регулярности с разбиением выборок наблюдений на обучающие и проверочные для моделирования в классе линейных регрессионных уравнений со случайными коэффициентами. Доказано существование оптимального множества регрессоров и выявлено условие редукции оптимальной регрессионной модели, зависимое от параметров модели и объемов выборок.

The regularity criterion with dividing of observation sample on training and testing samples for modeling in a class of linear regression equations with the random coefficients is researched. The existing of the optimum regressors set is proved and the condition of the optimal regression model reduction is obtained. This condition depends on parameters of model and volumes of samples.

Досліджено критерій регулярності з розбиттям вибірок спостережень на навчальні й перевірні вибірки для моделювання в класі лінійних регресійних рівнянь з випадковими коефіцієнтами. Доведено існування оптимальної множини регресорів та виявлено умову редукції оптимальної регресійної моделі, яка залежить від параметрів моделі та обсягів вибірок.

Введение. Задачи моделирования по результатам наблюдений в классе регрессионных уравнений часто бывают поставлены в условиях структурной неопределенности по количеству и составу входных переменных, и для их решения необходимо принять критерий для оценивания качества и сравнения моделей с разными структурами. Известный подход к построению критериев качества в условиях структурной неопределенности применяется в методе группового учета аргументов (МГУА), разработанный академиком НАН Украины А.Г. Ивахненко [1–8]. Подход основан на разбиении выборки данных на обучающую и проверочную части: на обучающей выборке оцениваются коэффициенты моделей, а на проверочной оценивается качество моделей.

Распространенный класс моделей – класс линейных по параметрам регрессионных уравнений со случайными коэффициентами. Модели этого класса позволяют описывать и прогнозировать состояния объектов как линейных по входным переменным, так и нелинейных (необходимо предварительно расширить множество входных переменных за счет нелинейных функций), а также применять и при моделировании статических характеристик динамических систем. В сравнении с классом обычных регрессионных уравнений этому классу моделей в МГУА уделено недостаточно внимания. Поэтому исследование критерия регулярности для структурной идентификации в классе регрессионных моделей со случайными коэффициентами – актуальная задача.

Постановка задачи

Пусть модель статического объекта представляет собой регрессионное уравнение

$$y_i = \overset{\circ}{\mathbf{x}}{}_i^{\mathrm{T}}\, \boldsymbol{\theta}_i + \zeta_i, \quad i = 1, 2, \ldots, n, \tag{1}$$

где y_i – i-е наблюдение выходной переменной; $\overset{\circ}{\mathbf{x}}_i$ – $(\overset{\circ}{m}\times 1)$-вектор i-го наблюдения множества входов $\overset{\circ}{X} = \{x_1, x_2, \ldots, x_{\overset{\circ}{m}}\}$ ($\overset{\circ}{X} \neq \varnothing$, $\varnothing$ – пустое множество), которые участвуют в формировании выходной переменной y; $\boldsymbol{\theta}_i$ – ненаблюдаемый случайный $(\overset{\circ}{m}\times 1)$-вектор коэффициентов; ξ_i – ненаблюдаемая случайная величина; n – объем выборки наблюдений.

Пусть для случайного вектора $\boldsymbol{\theta}_i$ выполнено

$$\boldsymbol{\theta}_i = \overset{\circ}{\boldsymbol{\theta}} + \boldsymbol{\eta}_i, \quad i = 1, 2, \ldots, n, \tag{2}$$

где $\overset{\circ}{\boldsymbol{\theta}} = (\overset{\circ}{\theta}_1, \overset{\circ}{\theta}_2, \ldots, \overset{\circ}{\theta}_{\overset{\circ}{m}})^{\mathrm{T}}$ – неизвестный детерминированный $(\overset{\circ}{m}\times 1)$-вектор; $\boldsymbol{\eta}_i$ – случайный $(\overset{\circ}{m}\times 1)$-вектор.

Пусть вектор $\boldsymbol{\eta}_i = (\eta_{1i}, \eta_{2i}, \ldots, \eta_{\overset{\circ}{m},i})^{\mathrm{T}}$ распределен по $\overset{\circ}{m}$-мерному нормальному закону: $\boldsymbol{\eta}_i \sim N(\mathbf{0}_{\overset{\circ}{m}}, \boldsymbol{\Sigma}_\eta)$, а для случайных векторов $\boldsymbol{\eta}_i$ выполнено

$$E\{\boldsymbol{\eta}_i\} = \mathbf{0}_{\overset{\circ}{m}}, \quad i = 1,2...,n; \tag{3}$$

$$E\{\boldsymbol{\eta}_i \boldsymbol{\eta}_i^{\mathrm{T}}\} = \boldsymbol{\Sigma}_{\eta}(\overset{\circ}{X}); \tag{4}$$

$$E\{\boldsymbol{\eta}_{i_1} \boldsymbol{\eta}_{i_2}^{\mathrm{T}}\} = \mathbf{O}_{(\overset{\circ}{m}\times\overset{\circ}{m})}, \; i_1, i_2 = 1,2,...,n, \; i_1 \neq i_2, \tag{5}$$

где $E\{\cdot\}$ – знак математического ожидания по реализациям случайных векторов $\boldsymbol{\eta}_i$; $\mathbf{0}_{\overset{\circ}{m}}$ – нулевой $(\overset{\circ}{m}\times 1)$-вектор; $\boldsymbol{\Sigma}_{\eta}(\overset{\circ}{X})$ – ковариационная $(\overset{\circ}{m}\times\overset{\circ}{m})$-матрица.

Пусть относительно $\boldsymbol{\zeta}$ выполнено

$$E\{\boldsymbol{\zeta}\} = \mathbf{0}_n, \quad E\{\boldsymbol{\zeta}\boldsymbol{\zeta}^{\mathrm{T}}\} = \sigma_{\zeta}\,\mathbf{I}_n, \tag{6}$$

$$E\{\zeta_{i_1}\zeta_{i_2}\} = 0, \quad i_1, i_2 = 1, \dots, n, \quad i_1 \neq i_2, \tag{7}$$

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

Предположим, что флуктуации коэффициентов $\boldsymbol{\eta}_i$ в (2) и аддитивная составляющая ζ_i в (1) статистически независимы

$$E\{\boldsymbol{\eta}_i\,\zeta_i\} = \mathbf{0}_{\overset{\circ}{m}}, \quad i = 1,2...,n. \tag{8}$$

Пусть в результате наблюдения объекта получены: 1) $\overset{\circ}{\mathbf{X}}$ – $(n\times\overset{\circ}{m})$-матрица n наблюдений $\overset{\circ}{m}$ входов множества $\overset{\circ}{X}$, имеющая полный ранг, равный $\overset{\circ}{m}$; 2) $\mathbf{y}$ – $(n\times 1)$-вектор наблюдений выходной переменной y.

В соответствии с (1)–(2) для наблюдений выполняется

$$\begin{aligned} y_i &= \overset{\circ}{\mathbf{x}}{}_i^{\mathrm{T}}\overset{\circ}{\boldsymbol{\theta}} + \overset{\circ}{\mathbf{x}}{}_i^{\mathrm{T}}\boldsymbol{\eta}_i + \zeta_i = \\ &= \overset{\circ}{\mathbf{x}}{}_i^{\mathrm{T}}\overset{\circ}{\boldsymbol{\theta}} + \xi_i(\overset{\circ}{X}) = \overset{\circ}{y}_i + \xi_i(\overset{\circ}{X}), \end{aligned} \tag{9}$$

$$\xi_i(\overset{\circ}{X}) = \overset{\circ}{\mathbf{x}}{}_i^{\mathrm{T}}\boldsymbol{\eta}_i + \zeta_i, \quad i = 1,2...,n. \tag{10}$$

Для математического ожидания ξ_i, учитывая (3) и (6), получаем

$$E\{\xi_i(\overset{\circ}{X})\} = 0, \quad i = 1,2,\dots,n, \tag{11}$$

а для дисперсий и ковариаций случайных величин $\xi_i(\overset{\circ}{X})$ с учетом (4)–(5) и (6)–(7) получаем

$$\begin{aligned} [\boldsymbol{\Sigma}_{\xi}(\overset{\circ}{X})]_{ii} &= E\{\xi_i(\overset{\circ}{X})\xi_i(\overset{\circ}{X})\} = \\ &= \overset{\circ}{\mathbf{x}}{}_i^{\mathrm{T}}\boldsymbol{\Sigma}_{\eta}(\overset{\circ}{X})\overset{\circ}{\mathbf{x}}_i + \sigma_{\zeta} = [\boldsymbol{\Lambda}_{\eta}(\overset{\circ}{X})]_{ii} + \sigma_{\zeta}, \end{aligned} \tag{12}$$

$$\boldsymbol{\Sigma}_{\xi}(\overset{\circ}{X}) = \boldsymbol{\Lambda}_{\eta}(\overset{\circ}{X}) + \sigma_{\zeta}\cdot\mathbf{I}_n, \tag{13}$$

$$E\{\xi_{i_1}(\overset{\circ}{X})\xi_{i_2}(\overset{\circ}{X})\} = 0, \; i_1, i_2 = 1, \dots, n, \; i_1 \neq i_2, \tag{14}$$

где $\boldsymbol{\Lambda}_{\eta}(\overset{\circ}{X})$ – диагональная $(n\times n)$-матрица, ее элементы определены в (12).

Для математического ожидания $(n\times 1)$-вектора $\boldsymbol{\xi}(\overset{\circ}{X}) = (\xi_1(\overset{\circ}{X}), \xi_2(\overset{\circ}{X}), \dots, \xi_n(\overset{\circ}{X}))^{\mathrm{T}}$ имеем

$$E\{\boldsymbol{\xi}^{\mathrm{T}}(\overset{\circ}{X})\boldsymbol{\xi}(\overset{\circ}{X})\} = \sum_{i=1}^{n}\overset{\circ}{\mathbf{x}}{}_i^{\mathrm{T}}\boldsymbol{\Sigma}_{\eta}(\overset{\circ}{X})\overset{\circ}{\mathbf{x}}_i + n\cdot\sigma_{\zeta}. \tag{15}$$

Формулы для оценивания коэффициентов

Запишем (9)–(15) в объединенном виде

$$\mathbf{y} = \overset{\circ}{\mathbf{X}}\overset{\circ}{\boldsymbol{\theta}} + \boldsymbol{\xi}(\overset{\circ}{X}) = \overset{\circ}{\mathbf{y}} + \boldsymbol{\xi}(\overset{\circ}{X}), \tag{16}$$

где $\mathbf{y}$ – $(n\times 1)$-вектор наблюдаемых зашумленных значений; $\overset{\circ}{\mathbf{X}}$ –$(n\times\overset{\circ}{m})$-матрица регрессоров множества $\overset{\circ}{X}$; $\overset{\circ}{\boldsymbol{\theta}}$ – неизвестный детерминированный $(\overset{\circ}{m}\times 1)$-вектор; $\overset{\circ}{\mathbf{y}}$ – $(n\times 1)$- вектор ненаблюдаемых значений; $\boldsymbol{\xi}(\overset{\circ}{X})$ – $(n\times 1)$- вектор ненаблюдаемых случайных величин в (10).

Необходимо найти оценку неизвестных коэффициентов $\overset{\circ}{\boldsymbol{\theta}}$ в виде

$$\mathbf{d} = \mathbf{C}\;\mathbf{y}, \tag{17}$$

где $(\overset{\circ}{m}\times n)$-матрицу $\mathbf{C}$, зависящую от $\overset{\circ}{\mathbf{X}}$, требуется определить.

Будем искать такую матрицу $\mathbf{C}$, при которой логарифм определителя ковариационной матрицы оценок коэффициентов (17) принимает минимальное значение и оценки коэффициентов несмещены [9]. Математическое ожидание и ковариационную матрицу оценки (17) вычислим по всем возможным реализациям случайных величин ξ. Для математического ожидания оценки (17) должно выполняться ($E\{\cdot\}$ – операция математического ожидания)

$$E\{\mathbf{d}\} = E\{\mathbf{C}\,\mathbf{y}\} = E\{\mathbf{C}\,(\overset{\circ}{\mathbf{y}} + \boldsymbol{\xi}(\overset{\circ}{X}))\} = \\ = E\{\mathbf{C}\,\overset{\circ}{\mathbf{X}}\overset{\circ}{\boldsymbol{\theta}}\} + E\{\mathbf{C}\,\boldsymbol{\xi}(\overset{\circ}{X})\} = \overset{\circ}{\boldsymbol{\theta}}. \tag{18}$$

Справедливость (18) следует из условий

$$\mathbf{C}\,\overset{\circ}{\mathbf{X}} = \mathbf{I}_m, \quad E\{\mathbf{C}\,\boldsymbol{\xi}(\overset{\circ}{X})\} = \mathbf{0}_m, \tag{19}$$

т.е. из несмещенности оценок и независимости элементов матрицы $\overset{\circ}{\mathbf{X}}$ от вектора $\xi(\overset{\circ}{X})$.

С учетом (18) для ковариационной матрицы оценок (17) выполняется

$$\mathrm{Cov}(\mathbf{d}) = E\{(\mathbf{d} - \overset{\circ}{\boldsymbol{\theta}})(\mathbf{d} - \overset{\circ}{\boldsymbol{\theta}})^{\mathrm{T}}\} = \\ E\{(\mathbf{C}\,\overset{\circ}{\mathbf{y}} + \mathbf{C}\boldsymbol{\xi}(\overset{\circ}{X}) - \overset{\circ}{\boldsymbol{\theta}})(\mathbf{C}\,\overset{\circ}{\mathbf{y}} + \mathbf{C}\boldsymbol{\xi}(\overset{\circ}{X}) - \overset{\circ}{\boldsymbol{\theta}})^{\mathrm{T}}\} = \\ = E\{\mathbf{C}\,\boldsymbol{\xi}(\overset{\circ}{X})\boldsymbol{\xi}^{\mathrm{T}}(\overset{\circ}{X})\mathbf{C}^{\mathrm{T}}\} = \mathbf{C}\,\boldsymbol{\Sigma}_{\xi}(\overset{\circ}{X})\mathbf{C}^{\mathrm{T}}. \tag{20}$$

Запишем функцию Лагранжа

$$L(\mathbf{C}, \boldsymbol{\Lambda}) = \ln\,(\det\,[\mathbf{C}\,\boldsymbol{\Sigma}_{\xi}(\overset{\circ}{X})\mathbf{C}^{\mathrm{T}}]) + \\ + \mathrm{tr}\,[\boldsymbol{\Lambda}_L(\mathbf{C}\,\overset{\circ}{\mathbf{X}} - \mathbf{I}_{\overset{\circ}{M}})], \tag{21}$$

где $\boldsymbol{\Lambda}_L$ – диагональная $-(\overset{\circ}{m} \times \overset{\circ}{m})$ матрица неопределенных множителей.

Тогда необходимые условия оптимальности имеют вид:

$$\begin{cases} \dfrac{\partial L}{\partial \mathbf{C}} = \dfrac{\partial}{\partial \mathbf{C}}\,(\ln\,(\det[\mathbf{C}\,\boldsymbol{\Sigma}_{\xi}(\overset{\circ}{X})\mathbf{C}^{\mathrm{T}}])) + \\ \qquad + \dfrac{\partial}{\partial \mathbf{C}}\,(\mathrm{tr}[\boldsymbol{\Lambda}_L(\mathbf{C}\,\overset{\circ}{\mathbf{X}} - \mathbf{I}_{\overset{\circ}{m}})]) = \mathbf{O}_{\overset{\circ}{m} \times m}, \\ \dfrac{\partial L}{\partial \boldsymbol{\Lambda}} = \dfrac{\partial}{\partial \boldsymbol{\Lambda}}\,(\mathrm{tr}[\boldsymbol{\Lambda}_L(\mathbf{C}\,\overset{\circ}{\mathbf{X}} - \mathbf{I}_{\overset{\circ}{m}})]) = \\ \qquad = \mathbf{C}\,\overset{\circ}{\mathbf{X}} - \mathbf{I}_{\overset{\circ}{m}} = \mathbf{O}_{\overset{\circ}{m} \times \overset{\circ}{m}}. \end{cases} \tag{22}$$

Применяя правила матричного дифференцирования, из (22) получаем

$$\mathbf{C} = (\overset{\circ}{\mathbf{X}}{}^{\mathrm{T}}\,\boldsymbol{\Sigma}_{\xi}^{-1}\,\overset{\circ}{\mathbf{X}})^{-1}\,\overset{\circ}{\mathbf{X}}{}^{\mathrm{T}}\,\boldsymbol{\Sigma}_{\xi}^{-1}. \tag{23}$$

Для оценки (17)–(23) выполняется

$$E\{\mathbf{d}\} = E\{\mathbf{C}\mathbf{y}\} = \\ = E\{(\overset{\circ}{\mathbf{X}}{}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}\,\overset{\circ}{\mathbf{X}})^{-1}\,\overset{\circ}{\mathbf{X}}{}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}\,(\overset{\circ}{\mathbf{X}}\overset{\circ}{\boldsymbol{\theta}} + \boldsymbol{\xi})\} = \overset{\circ}{\boldsymbol{\theta}}, \tag{24}$$

$$\mathrm{Cov}\,(\mathbf{d}) = E\{(\overset{\circ}{\mathbf{X}}{}^{\mathrm{T}}\,\boldsymbol{\Sigma}_{\xi}^{-1}\,\overset{\circ}{\mathbf{X}})^{-1}\,\overset{\circ}{\mathbf{X}}{}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}\,\boldsymbol{\xi}\boldsymbol{\xi}^{\mathrm{T}} \times \\ \times \boldsymbol{\Sigma}_{\xi}^{-1}\,\overset{\circ}{\mathbf{X}}(\overset{\circ}{\mathbf{X}}{}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}\,\overset{\circ}{\mathbf{X}})^{-1}\} = (\overset{\circ}{\mathbf{X}}{}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}\,\overset{\circ}{\mathbf{X}})^{-1}. \tag{25}$$

Критерий регулярности МГУА

В предыдущих разделах предполагалось, что множество регрессоров $\overset{\circ}{X}$, участвующих в формировании выходной переменной в (1), задано. Далее будем предполагать, что оно неизвестно и его требуется определить, т. е. рассмотрим задачу структурной идентификации. Для решения задачи структурной идентификации необходимо:

- указать метод оценивания коэффициентов в моделях с заданной структурой;
- построить алгоритм генерации различных структур моделей;
- принять способ сравнения моделей с разной структурой.

В формулах (17)–(25) предполагается, что ковариационная матрица $\boldsymbol{\Sigma}_{\eta}(\overset{\circ}{X})$ и скаляр σ_{ζ} априорно известны (случай, когда они неизвестны, будет предметом исследования отдельной работы). В условиях неопределенности по составу регрессоров класс моделей (1)–(2) можно трактовать таким образом, что все входные переменные из множества X могут влиять на выходную переменную, но математические ожидания коэффициентов влияния отличны от нуля только для подмножества регрессоров $\overset{\circ}{X} \subseteq X$. В данном разделе предположим, что известны $\boldsymbol{\Sigma}_{\eta}(X)$ – ковариационная матрица для множества X и скаляр σ_{ζ}.

В качестве алгоритма генерации перебираемых структур примем алгоритм полного перебора всех возможных структур; на его этапе с номером s в модели допускается только $s \le p$ регрессоров, где p – заданное максимально возможное число регрессоров в модели. Отметим, что в рассматриваемом классе регрессионных моделей (линейных по входам и коэффициентам) *структура* модели однозначно определяется составом множества входов, присутствующих в модели, а *сложность* модели – их числом.

Пусть X – заданное исходное множество m наблюдаемых входов, $\mathbf{X}$ – соответствующая $(n \times m)$-матрица регрессоров; $\overset{o}{X} \subseteq X$ – истинное множество $\overset{o}{m}$ входов, участвующих в модели (1), $\overset{o}{\mathbf{X}}$ – соответствующая $(n \times \overset{o}{m})$-матрица регрессоров истинного множества входов; $V \subseteq X$ – текущее множество s входов, которое меняется в ходе генерации различных структур, $\mathbf{V}$ – $(n \times s)$-матрица наблюдений входов, принадлежащих текущему множеству входов.

Рассмотрим функционал качества регрессионной модели, отражающий требование минимизации математического ожидания квадратичной нормы, называемый в литературе *J*-функционалом (см., например, [9]; в рамках МГУА этот функционал получил название *идеальный внешний критерий* [3]):

$$J(V, \mathbf{V}) = E\{\| \overset{o}{\mathbf{y}} - \hat{\mathbf{y}}(V, \mathbf{V}) \|^2\}, \qquad (26)$$

где $\mathbf{V}$ – $(n \times s)$-матрица наблюдений анализируемого множества входов V; $\hat{\mathbf{y}}(V, \mathbf{V}) = \mathbf{V}\hat{\mathbf{d}}$ – выход регрессионной модели, построенной на множестве входов V, а $\hat{\mathbf{d}}$ – оценка $(s\times 1)$-вектора коэффициентов регрессии y по V

$$\hat{\mathbf{d}} = (\mathbf{V}^{\mathrm{T}} \mathbf{\Sigma}_{\xi}^{-1}(V) \mathbf{V})^{-1} \mathbf{V}^{\mathrm{T}} \mathbf{\Sigma}_{\xi}^{-1}(V) \mathbf{y}, \qquad (27)$$

где $\mathbf{\Sigma}_{\xi}(V)$ – ковариационная матрица для текущего множества V:

$$\begin{aligned}[\mathbf{\Sigma}_{\xi}(V)]_{ii} &= E\{\xi_i(V)\xi_i(V)\} = \\ &= \mathbf{v}_i^{\mathrm{T}} \mathbf{\Sigma}_{\eta}(V) \mathbf{v}_i + \sigma_{\zeta} = [\mathbf{\Lambda}_{\eta}(V)]_{ii} + \sigma_{\zeta},\end{aligned} \qquad (28)$$

$$\mathbf{\Sigma}_{\xi}(V) = \mathbf{\Lambda}_{\eta}(V) + \sigma_{\zeta} \cdot \mathbf{I}_n, \qquad (29)$$

где $\mathbf{v}_i$ – $(s \times 1)$-вектор *i*-го наблюдения текущего множества входов V.

Определение. *J*-оптимальным множеством регрессоров называется такое множество входов $V_J \subseteq X$, для которого выполняется

$$V_J = \arg \min_{V \subseteq X} J(V, \mathbf{V}). \qquad (30)$$

• *J*-оптимальной по количеству и составу регрессоров называется регрессионная модель, построенная на множестве регрессоров V_J.

Как известно [3, 5, 8], *J*-оптимальное множество регрессоров может включать в себя не все регрессоры, соответствующие множеству входов $\overset{o}{X}$. В этом случае говорят о редукции (упрощении) *J*-оптимальной модели.

Функционал (30) не может применяться при решении практических задач (содержит ненаблюдаемый вектор $\overset{o}{\mathbf{y}}$), но может быть использован для теоретического сравнения методов оценивания, в том числе на основе метода статистических испытаний [8]. Если в качестве оценки $\overset{o}{\mathbf{y}}$ взять наблюдаемый вектор $\mathbf{y}$, то получаем так называемую *остаточную сумму квадратов* [10]: $RSS(\mathbf{V}) = \| \mathbf{y} - \hat{\mathbf{y}} \|^2$. Известно [3, 5, 8], что $E\{RSS(\mathbf{V})\}$ уменьшается с добавлением любого регрессора, даже заведомо избыточного, т.е. остаточная сумма квадратов не может быть использована для поиска *J*-оптимального множества регрессоров.

Существует ли альтернатива выбору ненаблюдаемого вектора $\overset{o}{\mathbf{y}}$ в *J*-функционале (30), обеспечивающая конструктивность соответствующего функционала качества регрессионной модели и сохраняющая для него свойства *J*-функционала? Для ответа на этот вопрос обратимся к критерию регулярности МГУА в схеме повторных наблюдений [8].

Пусть в условиях активного эксперимента имеется возможность для заданного *m*-мерного входа объекта делать не одно, а два независимых наблюдения выхода объекта. Будем относить первое наблюдение из пары наблюдений к обучающей выборке A, а второе наблюдение – к проверочной выборке B.

Пусть результатами наблюдения в условиях, когда разбиение осуществлено по схеме повторных наблюдений, есть *n*-мерные векторы

$\mathbf{y}_A$, $\mathbf{y}_B$ и матрица всех регрессоров $\mathbf{X}$, имеющая полный ранг, т.е. $\operatorname{rank}\mathbf{X}=m$, причем один из столбцов матрицы $\mathbf{X}$ состоит из единиц. Согласно (16) и принятому разбиению по схеме повторных независимых наблюдений для выборок A и B выполняется

$$\begin{aligned}\mathbf{y}_A &= \overset{o}{\mathbf{X}}\overset{o}{\boldsymbol{\theta}}+\boldsymbol{\xi}_A(\overset{o}{X})=\overset{o}{\mathbf{y}}+\boldsymbol{\xi}_A(\overset{o}{X}),\\ \mathbf{y}_B &= \overset{o}{\mathbf{X}}\overset{o}{\boldsymbol{\theta}}+\boldsymbol{\xi}_B(\overset{o}{X})=\overset{o}{\mathbf{y}}+\boldsymbol{\xi}_B(\overset{o}{X}),\end{aligned}\tag{31}$$

где $\boldsymbol{\xi}_A(\overset{o}{X})$ и $\boldsymbol{\xi}_B(\overset{o}{X})$ – $(n\times 1)$-векторы случайных величин, для которых выполнено

$$\begin{gathered}E\{\boldsymbol{\xi}_A(\overset{o}{X})\}=E\{\boldsymbol{\xi}_B(\overset{o}{X})\}=\mathbf{0}_n,\\ E\{\boldsymbol{\xi}_A(\overset{o}{X})\boldsymbol{\xi}_B^{\mathrm{T}}(\overset{o}{X})\}=\mathbf{O}_{n\times n},\end{gathered}\tag{32}$$

$$\begin{gathered}[E\{\boldsymbol{\xi}_A(\overset{o}{X})\boldsymbol{\xi}_A^{\mathrm{T}}(\overset{o}{X})\}]_{ii}=\\ =[E\{\boldsymbol{\xi}_B(\overset{o}{X})\boldsymbol{\xi}_B^{\mathrm{T}}(\overset{o}{X})\}]_{ii}=[\boldsymbol{\Sigma}_\xi(\overset{o}{X})]_{ii}=\\ =\overset{o}{\mathbf{x}}{}_i^{\mathrm{T}}\,\boldsymbol{\Sigma}_\eta(\overset{o}{X})\overset{o}{\mathbf{x}}_i+\sigma_\zeta=\\ =[\boldsymbol{\Lambda}_\eta(\overset{o}{X})]_{ii}+\sigma_\zeta,\quad i=1,2...,n,\end{gathered}\tag{33}$$

$$\boldsymbol{\Sigma}_\xi(\overset{o}{X})=\boldsymbol{\Lambda}_\eta(\overset{o}{X})+\sigma_\zeta\cdot\mathbf{I}_n,\tag{34}$$

где $\mathbf{O}_{n\times n}$ – нулевая $(n\times n)$-матрица; $\boldsymbol{\Lambda}_\eta(\overset{o}{X})$ – диагональная $(n\times n)$-матрица с элементами (33).

На обучающей выборке будем оценивать параметры системы регрессионных уравнений с текущей анализируемой структурой, а на проверочной – оценивать качество этой модели.

Рассмотрим случайную величину, называемую в методе группового учета аргументов критерием регулярности [1, 3, 5], в условиях, когда разбиение на обучающую и проверочную выборки осуществлено по схеме повторных наблюдений (для отличия от традиционного способа разбиения, без повторения наблюдений, в обозначение критерия введена звездочка):

$$AR^*(\mathbf{V})=(\mathbf{y}_B-\mathbf{V}\hat{\mathbf{d}}_A)^{\mathrm{T}}\boldsymbol{\Sigma}_\xi^{-1}(V)(\mathbf{y}_B-\mathbf{V}\hat{\mathbf{d}}_A),\tag{35}$$

где $\boldsymbol{\Sigma}_\xi(V)$ – определенная в (29) ковариационная матрица для текущего множества входов V; $\hat{\mathbf{d}}_A$ – оценка $(s\times 1)$-вектора коэффициентов регрессии y по V на выборке A, полученная по формуле (27) взвешенным методом наименьших квадратов

$$\begin{aligned}\hat{\mathbf{d}}_A&=(\mathbf{V}^{\mathrm{T}}\boldsymbol{\Sigma}_\xi^{-1}(V)\mathbf{V})^{-1}\mathbf{V}^{\mathrm{T}}\boldsymbol{\Sigma}_\xi^{-1}(V)\mathbf{y}_A=\\ &=(\mathbf{V}^{\mathrm{T}}\boldsymbol{\Sigma}_\xi^{-1}(V)\mathbf{V})^{-1}\mathbf{V}^{\mathrm{T}}\boldsymbol{\Sigma}_\xi^{-1}(V)\overset{o}{\mathbf{X}}\,\overset{o}{\boldsymbol{\theta}}+\\ &+(\mathbf{V}^{\mathrm{T}}\boldsymbol{\Sigma}_\xi^{-1}(V)\mathbf{V})^{-1}\mathbf{V}^{\mathrm{T}}\boldsymbol{\Sigma}_\xi^{-1}(V)\boldsymbol{\xi}_A.\end{aligned}\tag{36}$$

С учетом (36) для вектора остатков в критерии регулярности выполняется

$$\begin{aligned}\mathbf{u}_B&=\mathbf{y}_B-\mathbf{V}\hat{\mathbf{d}}_A=\overset{o}{\mathbf{X}}\overset{o}{\boldsymbol{\theta}}+\boldsymbol{\xi}_B-\\ &-\mathbf{V}(\mathbf{V}^{\mathrm{T}}\boldsymbol{\Sigma}_\xi^{-1}(V)\mathbf{V})^{-1}\mathbf{V}^{\mathrm{T}}\boldsymbol{\Sigma}_\xi^{-1}(V)\overset{o}{\mathbf{X}}\,\overset{o}{\boldsymbol{\theta}}-\\ &-\mathbf{V}(\mathbf{V}^{\mathrm{T}}\boldsymbol{\Sigma}_\xi^{-1}(V)\mathbf{V})^{-1}\mathbf{V}^{\mathrm{T}}\boldsymbol{\Sigma}_\xi^{-1}(V)\boldsymbol{\xi}_A=\\ &=\mathbf{S}_V\,\overset{o}{\mathbf{X}}\,\overset{o}{\boldsymbol{\theta}}+\boldsymbol{\xi}_B-\mathbf{P}_V\boldsymbol{\xi}_A,\end{aligned}\tag{37}$$

где

$$\begin{aligned}\mathbf{S}_V&=\mathbf{I}_n-\mathbf{V}(\mathbf{V}^{\mathrm{T}}\boldsymbol{\Sigma}_\xi^{-1}(V)\mathbf{V})^{-1}\mathbf{V}^{\mathrm{T}}\boldsymbol{\Sigma}_\xi^{-1}(V),\\ \mathbf{P}_V&=\mathbf{V}(\mathbf{V}^{\mathrm{T}}\boldsymbol{\Sigma}_\xi^{-1}(V)\mathbf{V})^{-1}\mathbf{V}^{\mathrm{T}}\boldsymbol{\Sigma}_\xi^{-1}(V)\end{aligned}\;-\tag{38}$$

идемпотентные матрицы; $\mathbf{S}_V\,\overset{o}{\mathbf{X}}\,\overset{o}{\boldsymbol{\theta}}$ – так называемое смещение, обусловленное выбором текущего множества регрессоров V вместо истинного множества $\overset{o}{X}$.

Учитывая (37), выполнение $\mathbf{S}_V\mathbf{P}_V=\mathbf{O}_{n\times n}$, идемпотентность матриц $\mathbf{S}_V$ и $\mathbf{P}_V$, для критерия регулярности в схеме повторных наблюдений получаем

$$\begin{aligned}AR^*(\mathbf{V})&=\mathbf{u}_B^{\mathrm{T}}\boldsymbol{\Sigma}_\xi^{-1}(V)\mathbf{u}_B=\\ &=\overset{o}{\boldsymbol{\theta}}{}^{\mathrm{T}}\overset{o}{\mathbf{X}}{}^{\mathrm{T}}\mathbf{S}_V^{\mathrm{T}}\,\boldsymbol{\Sigma}_\xi^{-1}(V)\,\mathbf{S}_V\,\overset{o}{\mathbf{X}}\overset{o}{\boldsymbol{\theta}}+\\ &+2\overset{o}{\boldsymbol{\theta}}{}^{\mathrm{T}}\overset{o}{\mathbf{X}}{}^{\mathrm{T}}\mathbf{S}_V^{\mathrm{T}}\,\boldsymbol{\Sigma}_\xi^{-1}(V)(\boldsymbol{\xi}_B-\mathbf{P}_V\boldsymbol{\xi}_A)+\\ &+(\boldsymbol{\xi}_B-\mathbf{P}_V\boldsymbol{\xi}_A)^{\mathrm{T}}\,\boldsymbol{\Sigma}_\xi^{-1}(V)\,(\boldsymbol{\xi}_B-\mathbf{P}_V\boldsymbol{\xi}_A).\end{aligned}\tag{39}$$

Вычислим математическое ожидание критерия регулярности (39), используя предположения (28)–(29) и (32)–(34):

$$\begin{aligned}E\{AR^*(V)\}&=\overset{o}{\boldsymbol{\theta}}{}^{\mathrm{T}}\overset{o}{\mathbf{X}}{}^{\mathrm{T}}\boldsymbol{\Sigma}_\xi^{-1}(V)\mathbf{S}_V\,\overset{o}{\mathbf{X}}\,\overset{o}{\boldsymbol{\theta}}+\\ &+\operatorname{tr}[\boldsymbol{\Sigma}_\xi^{-1}(V)\,\boldsymbol{\Sigma}_\xi(\overset{o}{X})]+\operatorname{tr}[\boldsymbol{\Sigma}_\xi^{-1}(V)\,\mathbf{P}_V\boldsymbol{\Sigma}_\xi(\overset{o}{X})],\end{aligned}\tag{40}$$

где $\mathring{\boldsymbol{\theta}}^{\mathrm{T}}\mathring{\mathbf{X}}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}(V)\mathbf{S}_V\mathring{\mathbf{X}}\mathring{\boldsymbol{\theta}}=\mathring{\mathbf{y}}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}(V)\mathbf{S}_V\mathring{\mathbf{y}}$ – составляющая, обусловленная выбором текущего множества регрессоров V вместо истинного множества $\mathring{X}$.

Исследование критерия регулярности МГУА

Установим свойства критерия регулярности МГУА. С этой целью исследуем, как изменяется математическое ожидание критерия в зависимости от состава множества регрессоров. В случае истинной структуры для математического ожидания критерия регулярности в схеме повторных наблюдений, используя (32)–(34) и (40), получаем

$$E\{AR^*(\mathring{X})\}=\mathring{\boldsymbol{\theta}}^{\mathrm{T}}\mathring{\mathbf{X}}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}(\mathring{X})\times$$
$$\times[\mathbf{I}_n-\mathring{\mathbf{X}}(\mathring{\mathbf{X}}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}(\mathring{X})\mathring{\mathbf{X}})^{-1}\mathring{\mathbf{X}}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}(\mathring{X})]\,\mathring{\mathbf{X}}\mathring{\boldsymbol{\theta}}+$$
$$+\mathrm{tr}\,[\boldsymbol{\Sigma}_{\xi}^{-1}(\mathring{X})\boldsymbol{\Sigma}_{\xi}(\mathring{X})]+$$
$$+\,\mathrm{tr}\,[\boldsymbol{\Sigma}_{\xi}^{-1}(\mathring{X})\mathring{\mathbf{X}}(\mathring{\mathbf{X}}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}(\mathring{X})\mathring{\mathbf{X}})^{-1}\mathring{\mathbf{X}}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}(\mathring{X})\boldsymbol{\Sigma}_{\xi}(\mathring{X})]=$$
$$=\mathring{\boldsymbol{\theta}}^{\mathrm{T}}\mathring{\mathbf{X}}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}(\mathring{X})\mathring{\mathbf{X}}\mathring{\boldsymbol{\theta}}-\mathring{\boldsymbol{\theta}}^{\mathrm{T}}\mathring{\mathbf{X}}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}(\mathring{X})\times$$
$$\times\mathring{\mathbf{X}}(\mathring{\mathbf{X}}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}(\mathring{X})\mathring{\mathbf{X}})^{-1}\mathring{\mathbf{X}}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}(\mathring{X})\mathring{\mathbf{X}}\mathring{\boldsymbol{\theta}}+$$
$$+\mathrm{tr}\,[\mathbf{I}_n]+\mathrm{tr}\,[\boldsymbol{\Sigma}_{\xi}^{-1}(\mathring{X})\times$$
$$\times\mathring{\mathbf{X}}(\mathring{\mathbf{X}}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}(\mathring{X})\mathring{\mathbf{X}})^{-1}\mathring{\mathbf{X}}^{\mathrm{T}}]=n+\mathring{m}. \qquad (41)$$

Случай недостающего регрессора. Рассмотрим случай, когда в модель ошибочно не включен один регрессор, хотя он участвует в формировании значения выходной переменной, и для простоты считаем, что это регрессор с номером $\mathring{m}$ из множества $\mathring{X}$. Тогда для истинного и текущего множества регрессоров и их матриц наблюдений выполняется

$$\mathring{X}=V\cup\mathring{x}(\mathring{m}),\ \mathring{\mathbf{X}}=[\,\mathbf{V}\mid\mathbf{m}\,],\ \mathring{\mathbf{X}}^{\mathrm{T}}=\begin{bmatrix}\mathbf{V}^{\mathrm{T}}\\ \hline \mathbf{m}^{\mathrm{T}}\end{bmatrix}, \qquad (42)$$

где $\mathring{x}(\mathring{m})$ – пропущенный вход, а $\mathbf{m}$ – соответствующий ему регрессор.

Введем обозначение

$$\mathring{\boldsymbol{\theta}}=\begin{pmatrix}\mathring{\boldsymbol{\theta}}(V)\\ \hline \mathring{\theta}(\mathring{m})\end{pmatrix},\ \ \mathring{\boldsymbol{\theta}}^{\mathrm{T}}=(\mathring{\boldsymbol{\theta}}^{\mathrm{T}}(V),\mathring{\theta}(\mathring{m}))\,. \qquad (43)$$

В (40) для составляющей, обусловленной ошибкой в выборе структуры (42), выполняется

$$\mathring{\boldsymbol{\theta}}^{\mathrm{T}}\mathring{\mathbf{X}}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}(V)\mathbf{S}_V\mathring{\mathbf{X}}\mathring{\boldsymbol{\theta}}=\mathring{\boldsymbol{\theta}}^{\mathrm{T}}\mathring{\mathbf{X}}^{\mathrm{T}}\times$$
$$\times\Big[\,\boldsymbol{\Sigma}_{\xi}^{-1}(V)-\boldsymbol{\Sigma}_{\xi}^{-1}(V)\mathbf{V}(\mathbf{V}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}(V)\mathbf{V})^{-1}\mathbf{V}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}(V)\,\Big]\mathring{\mathbf{X}}\mathring{\boldsymbol{\theta}}=$$
$$=\mathring{\boldsymbol{\theta}}^{\mathrm{T}}\begin{bmatrix}\mathbf{V}^{\mathrm{T}}\\ \hline \mathbf{m}^{\mathrm{T}}\end{bmatrix}\Big[\boldsymbol{\Sigma}_{\xi}^{-1}(V)-\boldsymbol{\Sigma}_{\xi}^{-1}(V)\times$$
$$\times\mathbf{V}(\mathbf{V}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}(V)\mathbf{V})^{-1}\mathbf{V}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}(V)\Big]\Big[\,\mathbf{V}\mid\mathbf{m}\,\Big]\mathring{\boldsymbol{\theta}}=$$
$$=\mathring{\boldsymbol{\theta}}^{\mathrm{T}}\begin{bmatrix}\mathbf{V}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}(V)-\mathbf{V}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}(V)\mathbf{V}(\mathbf{V}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}(V)\mathbf{V})^{-1}\mathbf{V}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}(V)\\ \hline \mathbf{m}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}(V)-\mathbf{m}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}(V)\mathbf{V}(\mathbf{V}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}(V)\mathbf{V})^{-1}\mathbf{V}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}(V)\end{bmatrix}\times$$
$$\times\Big[\,\mathbf{V}\mid\mathbf{m}\,\Big]\mathring{\boldsymbol{\theta}}=$$
$$=\mathring{\boldsymbol{\theta}}^{\mathrm{T}}\begin{bmatrix}\mathbf{O}_{(s\times n)}\\ \hline \mathbf{m}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}(V)-\mathbf{m}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}(V)\mathbf{V}(\mathbf{V}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}(V)\mathbf{V})^{-1}\mathbf{V}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}(V)\end{bmatrix}\times$$
$$\times\Big[\,\mathbf{V}\mid\mathbf{m}\,\Big]\mathring{\boldsymbol{\theta}}=(\mathring{\boldsymbol{\theta}}^{\mathrm{T}}(V),\mathring{\theta}(\mathring{m}))\times$$
$$\times\left[\begin{array}{c|c}\mathbf{O}_{(s\times s)} & \mathbf{0}_s\\ \hline \mathbf{0}_s^{\mathrm{T}} & \begin{array}{l}\mathbf{m}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}(V)\mathbf{m}-\\ -\mathbf{m}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}(V)\mathbf{V}(\mathbf{V}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}(V)\mathbf{V})^{-1}\mathbf{V}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}(V)\mathbf{m}\end{array}\end{array}\right]\begin{pmatrix}\mathring{\boldsymbol{\theta}}(V)\\ \hline \mathring{\theta}(\mathring{m})\end{pmatrix}=$$
$$=\mathring{\theta}(\mathring{m})\mathbf{m}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}(V)\times$$
$$\times[\,(\mathbf{I}_n-\mathbf{V}(\mathbf{V}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}(V)\mathbf{V})^{-1}\mathbf{V}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}(V))\,]\,\mathbf{m}\mathring{\theta}(\mathring{m})=$$
$$=\mathring{\theta}(\mathring{m})\mathbf{m}^{\mathrm{T}}\boldsymbol{\Sigma}_{\xi}^{-1}(V)\mathbf{S}_V\,\mathbf{m}\mathring{\theta}(\mathring{m})\,. \qquad (44)$$

Для вычисления следов двух матриц в (40) установим соотношение диагональных матриц $\boldsymbol{\Sigma}_{\xi}(\mathring{X})$ и $\boldsymbol{\Sigma}_{\xi}(V)$, используя (28)–(29) и (32)–(34):

$$[\,\boldsymbol{\Sigma}_{\xi}(\mathring{X})]_{ii}=\mathring{\mathbf{x}}_i^{\mathrm{T}}\,\boldsymbol{\Sigma}_{\eta}(\mathring{X})\,\mathring{\mathbf{x}}_i+\sigma_{\zeta}=[\,\boldsymbol{\Lambda}_{\eta}(\mathring{X})]_{ii}+\sigma_{\zeta}=$$
$$=\left(\mathbf{v}_i^{\mathrm{T}},\mathring{x}_i(\mathring{m})\right)\left[\begin{array}{c|c}\boldsymbol{\Sigma}_{\eta}(V) & \boldsymbol{\sigma}_{\eta}(V,\mathring{x})\\ \hline \boldsymbol{\sigma}_{\eta}^{\mathrm{T}}(V,\mathring{x}) & \sigma_{\eta}(\mathring{x},\mathring{x})\end{array}\right]\begin{pmatrix}\mathbf{v}_i\\ \hline \mathring{x}_i(\mathring{m})\end{pmatrix}+\sigma_{\zeta}=$$
$$=\Big(\,\mathbf{v}_i^{\mathrm{T}}\boldsymbol{\Sigma}_{\eta}(V)+\mathring{x}_i(\mathring{m})\boldsymbol{\sigma}_{\eta}^{\mathrm{T}}(V,\mathring{x}),\mathbf{v}_i^{\mathrm{T}}\boldsymbol{\sigma}_{\eta}(V,\mathring{x})+$$

$$+\mathring{x}_i(\mathring{m})\sigma_\eta(\mathring{x},\mathring{x})\left(\begin{array}{c}\mathbf{v}_i\\ \hdashline \mathring{x}_i(\mathring{m})\end{array}\right)+\sigma_\zeta =$$

$$=\mathbf{v}_i^{\mathrm{T}}\mathbf{\Sigma}_\eta(V)\mathbf{v}_i+\mathring{x}_i(\mathring{m})\boldsymbol{\sigma}_\eta^{\mathrm{T}}(V,\mathring{x})\mathbf{v}_i+$$

$$+\mathbf{v}_i^{\mathrm{T}}\boldsymbol{\sigma}_\eta(V,\mathring{x})\mathring{x}_i(\mathring{m})+\mathring{x}_i(\mathring{m})\sigma_\eta(\mathring{x},\mathring{x})\mathring{x}_i(\mathring{m})+\sigma_\zeta =$$

$$=[\mathbf{\Sigma}_\xi(V)]_{ii}+\delta_{ii}(\sigma),\ \ i=1,2...,n\,, \tag{45}$$

где

$$\delta_{ii}(\sigma)=\mathring{x}_i(\mathring{m})\boldsymbol{\sigma}_\eta^{\mathrm{T}}(V,\mathring{x})\mathbf{v}_i+$$
$$+\mathbf{v}_i^{\mathrm{T}}\boldsymbol{\sigma}_\eta(V,\mathring{x})\mathring{x}_i(\mathring{m})+\mathring{x}_i(\mathring{m})\sigma_\eta(\mathring{x},\mathring{x})\mathring{x}_i(\mathring{m}). \tag{46}$$

Итак, установлено, что для матриц $\mathbf{\Sigma}_\xi(\mathring{X})$ и $\mathbf{\Sigma}_\xi(V)$ выполняется

$$\mathbf{\Sigma}_\xi(\mathring{X})=\mathbf{\Sigma}_\xi(V)+\mathbf{\Delta}(\sigma)\,, \tag{47}$$

где $\mathbf{\Delta}(\sigma)=diag\ \{\delta_{11}(\sigma),\delta_{22}(\sigma),\ldots,\delta_{nn}(\sigma)\}$ – диагональная матрица.

Запишем следы двух матриц в (40), используя (45)–(47):

$$\begin{aligned}&\mathrm{tr}\,[\mathbf{\Sigma}_\xi^{-1}(V)\mathbf{\Sigma}_\xi(\mathring{X})]=\\&=\mathrm{tr}\,[\mathbf{\Sigma}_\xi^{-1}(V)\mathbf{\Sigma}_\xi(V)]+\mathrm{tr}\,[\mathbf{\Sigma}_\xi^{-1}(V)\mathbf{\Delta}(\sigma)]]=\\&=\mathrm{tr}\,[\mathbf{I}_n]+\mathrm{tr}\,[\mathbf{\Sigma}_\xi^{-1}(V)\mathbf{\Delta}(\sigma)]]=\\&=n+\mathrm{tr}\,[\mathbf{\Sigma}_\xi^{-1}(V)\mathbf{\Delta}(\sigma)]],\end{aligned} \tag{48}$$

$$\begin{aligned}&\mathrm{tr}\,[\mathbf{\Sigma}_\xi^{-1}(V)\mathbf{V}(\mathbf{V}^{\mathrm{T}}\mathbf{\Sigma}_\xi^{-1}(V)\mathbf{V})^{-1}\mathbf{V}^{\mathrm{T}}\mathbf{\Sigma}_\xi^{-1}(V)\mathbf{\Sigma}_\xi(\mathring{X})]=\\&=\mathrm{tr}\,[\mathbf{\Sigma}_\xi^{-1}(V)\mathbf{V}(\mathbf{V}^{\mathrm{T}}\mathbf{\Sigma}_\xi^{-1}(V)\mathbf{V})^{-1}\mathbf{V}^{\mathrm{T}}]+\\&+\mathrm{tr}\,[\mathbf{\Sigma}_\xi^{-1}(V)\mathbf{V}(\mathbf{V}^{\mathrm{T}}\mathbf{\Sigma}_\xi^{-1}(V)\mathbf{V})^{-1}\mathbf{V}^{\mathrm{T}}\mathbf{\Sigma}_\xi^{-1}(V)\mathbf{\Delta}(\sigma)]=\\&=\mathrm{tr}\,[\mathbf{I}_s]+\mathrm{tr}\,[\mathbf{\Sigma}_\xi^{-1}(V)\times\\&\times\mathbf{V}(\mathbf{V}^{\mathrm{T}}\mathbf{\Sigma}_\xi^{-1}(V)\mathbf{V})^{-1}\mathbf{V}^{\mathrm{T}}\mathbf{\Sigma}_\xi^{-1}(V)\mathbf{\Delta}(\sigma)]=\\&=s+\mathrm{tr}\,[\mathbf{\Sigma}_\xi^{-1}(V)\mathbf{P}_V\mathbf{\Delta}(\sigma)].\end{aligned} \tag{49}$$

Объединяя результаты (44), (48) и (49), запишем разность

$$\begin{aligned}&\Delta_1(V,\mathring{X})=E\{AR^*(V)\}-E\{AR^*(\mathring{X})\}=\\&=\mathring{\theta}(\mathring{m})\mathbf{m}^{\mathrm{T}}\mathbf{\Sigma}_\xi^{-1}(V)\mathbf{S}_V\,\mathbf{m}\mathring{\theta}(\mathring{m})+\\&+n+\mathrm{tr}\,[\mathbf{\Sigma}_\xi^{-1}(V)\mathbf{\Delta}(\sigma)])]+\\&+(\mathring{m}-1)+\mathrm{tr}\,[\mathbf{\Sigma}_\xi^{-1}(V)\mathbf{P}_V\mathbf{\Delta}(\sigma)]-n-\mathring{m}=\\&=\mathring{\theta}(\mathring{m})\mathbf{m}^{\mathrm{T}}\mathbf{\Sigma}_\xi^{-1}(V)\mathbf{S}_V\,\mathbf{m}\mathring{\theta}(\mathring{m})+\mathrm{tr}\,[\mathbf{\Sigma}_\xi^{-1}(V)\mathbf{\Delta}(\sigma)])]-\\&-1+\mathrm{tr}\,[\mathbf{\Sigma}_\xi^{-1}(V)\mathbf{P}_V\mathbf{\Delta}(\sigma)],\end{aligned} \tag{50}$$

где матрицы $\mathbf{S}_V$ и $\mathbf{P}_V$ введены в (38), а оба следа положительны ввиду положительной определенности матриц под знаком следа.

Если $\Delta_1(V,\mathring{X})>0$, то структура $\mathring{X}$ лучше V; если $\Delta_1(V,\mathring{X})<0$, то структура V лучше $\mathring{X}$; если $\Delta_1(V,\mathring{X})=0$, то структура V лучше $\mathring{X}$ по принципу простоты.

Выполнение $\Delta_1(V,\mathring{X})\le 0$ – условие так называемой редукции модели, оптимальной по структуре. Из (50) для условия редукции получаем

$$\begin{aligned}&\mathring{\theta}(\mathring{m})\mathbf{m}^{\mathrm{T}}\mathbf{\Sigma}_\xi^{-1}(V)\mathbf{S}_V\,\mathbf{m}\mathring{\theta}(\mathring{m})+\mathrm{tr}\,[\mathbf{\Sigma}_\xi^{-1}(V)\mathbf{\Delta}(\sigma)])]+\\&+\mathrm{tr}\,[\mathbf{\Sigma}_\xi^{-1}(V)\mathbf{V}(\mathbf{V}^{\mathrm{T}}\mathbf{\Sigma}_\xi^{-1}(V)\mathbf{V})^{-1}\times\\&\times\mathbf{V}^{\mathrm{T}}\mathbf{\Sigma}_\xi^{-1}(V)\mathbf{\Delta}(\sigma)]\le 1.\end{aligned} \tag{51}$$

Получить условие редукции (51) в простом виде удается только при выполнении дополнительных предположений, и это предмет отдельного исследования.

Косвенное подтверждение истинности условия редукции. Установим здесь, какой вид принимает условие (51), если предположить, что коэффициенты в уравнении (1) есть детерминированными, а не случайными (для такого класса моделей задача структурной идентификации в условиях повторных наблюдений рассмотрена в [8, 11, 12]):

$$\boldsymbol{\theta}_i(k)=\mathring{\boldsymbol{\theta}}(k),\ \ i=1,2,\ldots,n\,. \tag{52}$$

При выполнении (52) матрицы $\mathbf{\Lambda}_\eta(\mathring{X})$ в (13), $\mathbf{\Lambda}_\eta(V)$ в (29) и $\mathbf{\Delta}(\sigma)$ в (47) есть нулевыми, а для матриц $\mathbf{\Sigma}_\xi(\mathring{X})$ и $\mathbf{\Sigma}_\xi(V)$ выполняется

$$\mathbf{\Sigma}_\xi(\mathring{X})=\mathbf{\Sigma}_\xi(V)=\sigma_\zeta\cdot\mathbf{I}_n\,. \tag{53}$$

Тогда для условия редукции (51) при выполнении (52) получаем

$$(\overset{\circ}{\theta}(\overset{\circ}{m}))^2\,\mathbf{m}^{\mathrm{T}}\,\mathbf{S}_V\,\mathbf{m} \le \sigma_\zeta\,. \qquad (54)$$

Совпадение (54) с результатами [8, 11 12] служит косвенным подтверждением истинности условия редукции (51).

Редукция модели, оптимальной по составу регрессоров, означает, что при выполнении соотношения между параметрами модели (54) следует исключить регрессор $\mathbf{m}$ из модели. Редуцированная модель будет иметь меньшую ошибку прогнозирования выходной переменной на новых выборках наблюдений в сравнении с истинной моделью.

Из (54) следует, что возможность редукции модели может быть обусловлена пятью причинами:

а) малостью нормы коэффициента $\overset{\circ}{\theta}(\overset{\circ}{m})$;

б) малостью нормы вектора наблюдений регрессора $\mathbf{m}$;

в) малым объемом выборок наблюдений n;

г) высокой степенью линейной зависимости регрессора $\mathbf{m}$ с другими регрессорами в матрице $\mathbf{V}$;

д) большим значением дисперсии σ_ζ.

Случай избыточного регрессора. Рассмотрим случай, когда в модель ошибочно включен излишний регрессор, хотя он не участвует в формировании значения выходной переменной. Тогда для текущего и истинного множества регрессоров и их матриц наблюдений выполняется

$$V=\overset{\circ}{X}\cup r,\quad \mathbf{V}=\left[\begin{array}{c:c}\overset{\circ}{\mathbf{X}} & \mathbf{r}\end{array}\right],\quad \mathbf{V}^{\mathrm{T}}=\left[\begin{array}{c}\overset{\circ}{\mathbf{X}}{}^{\mathrm{T}}\\ \hdashline \mathbf{r}^{\mathrm{T}}\end{array}\right], \qquad (55)$$

где r – излишний вход, а $\mathbf{r}$ – соответствующий ему избыточный регрессор.

В этом случае введенная в (40) составляющая, обусловленная выбором текущего множества регрессоров V вместо истинного множества $\overset{\circ}{X}$, равна нулю. Действительно, учитывая (55), получаем

$$\overset{\circ}{\boldsymbol{\theta}}{}^{\mathrm{T}}\overset{\circ}{\mathbf{X}}{}^{\mathrm{T}}\boldsymbol{\Sigma}_\xi^{-1}(V)\mathbf{S}_V\,\overset{\circ}{\mathbf{X}}\,\overset{\circ}{\boldsymbol{\theta}} =$$

$$=(\sigma_\zeta)^{-1}\cdot\overset{\circ}{\boldsymbol{\theta}}{}^{\mathrm{T}}\overset{\circ}{\mathbf{X}}{}^{\mathrm{T}}\left[\,\mathbf{I}_n-\mathbf{V}(\mathbf{V}^{\mathrm{T}}\mathbf{V})^{-1}\mathbf{V}^{\mathrm{T}}\right]\overset{\circ}{\mathbf{X}}\,\overset{\circ}{\boldsymbol{\theta}} =$$

$$=(\sigma_\zeta)^{-1}\cdot\overset{\circ}{\boldsymbol{\theta}}{}^{\mathrm{T}}\overset{\circ}{\mathbf{X}}{}^{\mathrm{T}}\times$$

$$\times\left[\mathbf{I}_n-\left[\begin{array}{c:c}\overset{\circ}{\mathbf{X}} & \mathbf{r}\end{array}\right]\left(\left[\begin{array}{c}\overset{\circ}{\mathbf{X}}{}^{\mathrm{T}}\\ \hdashline \mathbf{r}^{\mathrm{T}}\end{array}\right]\left[\begin{array}{c:c}\overset{\circ}{\mathbf{X}} & \mathbf{r}\end{array}\right]\right)^{-1}\left[\begin{array}{c}\overset{\circ}{\mathbf{X}}{}^{\mathrm{T}}\\ \hdashline \mathbf{r}^{\mathrm{T}}\end{array}\right]\right]\overset{\circ}{\mathbf{X}}\,\overset{\circ}{\boldsymbol{\theta}} =$$

$$=(\sigma_\zeta)^{-1}\cdot\overset{\circ}{\boldsymbol{\theta}}{}^{\mathrm{T}}\overset{\circ}{\mathbf{X}}{}^{\mathrm{T}}\times$$

$$\times\left[\mathbf{I}_n-\left[\begin{array}{c:c}\overset{\circ}{\mathbf{X}} & \mathbf{r}\end{array}\right]\left(\left[\begin{array}{c:c}\overset{\circ}{\mathbf{X}}{}^{\mathrm{T}}\overset{\circ}{\mathbf{X}} & \overset{\circ}{\mathbf{X}}{}^{\mathrm{T}}\mathbf{r}\\ \hdashline \mathbf{r}^{\mathrm{T}}\overset{\circ}{\mathbf{X}} & \mathbf{r}^{\mathrm{T}}\mathbf{r}\end{array}\right]\right)^{-1}\left[\begin{array}{c}\overset{\circ}{\mathbf{X}}{}^{\mathrm{T}}\\ \hdashline \mathbf{r}^{\mathrm{T}}\end{array}\right]\right]\overset{\circ}{\mathbf{X}}\,\overset{\circ}{\boldsymbol{\theta}} =$$

$$=(\sigma_\zeta)^{-1}\cdot\overset{\circ}{\boldsymbol{\theta}}{}^{\mathrm{T}}\,\mathbf{O}_{(s\times s)}\,\overset{\circ}{\mathbf{X}}\overset{\circ}{\boldsymbol{\theta}}=0\,. \qquad (56)$$

Для перемножения блочных матриц в (56) применена формула обращения блочной матрицы (частный случай формулы Фробениуса [9]).

С учетом (56) для избыточного регрессора получаем

$$\Delta_2(V,\overset{\circ}{X})=E\{AR^*(V)\}-E\{AR^*(\overset{\circ}{X})\}=$$

$$=\mathrm{tr}[\boldsymbol{\Sigma}_\xi^{-1}(V)\boldsymbol{\Sigma}_\xi(\overset{\circ}{X})]+\mathrm{tr}[\boldsymbol{\Sigma}_\xi^{-1}(V)\mathbf{P}_V\boldsymbol{\Sigma}_\xi(\overset{\circ}{X})]-n-\overset{\circ}{m}=$$

$$=\mathrm{tr}\,[\,\mathbf{I}_n]+\mathrm{tr}\,[\,\mathbf{I}_{(\overset{\circ}{m}+1)}]-n-\overset{\circ}{m}=1\,. \qquad (57)$$

Из (57) следует, что в случае избыточного регрессора истинная структура всегда лучше, а регрессор $\mathbf{r}$ действительно не следует включать в модель.

Заключение. В соответствии с принципами метода группового учета аргументов построен и исследован критерий структурной идентификации для моделирования в классе регрессионных уравнений со случайными коэффициентами. В схеме повторных наблюдений получено условие редукции (упрощения) оптимального по составу регрессоров регрессионного уравнения. Условие зависит от параметров модели и объемов выборок. В частном случае, когда коэффициенты есть детерминированными величинами, это условие совпадает с известным в МГУА условием редукции обычного регрессионного уравнения.

Окончание на стр. 29

1. *Ивахненко А.Г.* Индуктивный метод самоорганизации моделей сложных систем. – Киев: Наук. думка, 1982. – 296 с.
2. *Self-organizing* methods in modelling: GMDH type algorithms / Ed. by S.J. Farlow. – New York, Basel: Marcel Decker Inc., 1984. – 350 p.
3. *Ивахненко А.Г.*, *Степашко В.С.* Помехоустойчивость моделирования. – Киев: Наук. думка, 1985. – 216 с.
4. *Ивахненко А.Г.*, *Мюллер Й.А.* Самоорганизация прогнозирующих моделей. – Киев: Техніка, 1985. – 223 с.
5. *Ивахненко А.Г.*, *Юрачковский Ю.П.* Моделирование сложных систем по экспериментальным данным. – М. : Радио и связь, 1987. – 120 с.
6. *Madala H.R.*, *Ivakhnenko A.G.* Inductive Learning Algorithms for Complex System Modeling. – London, Tokyo: CRC Press Inc., 1994. – 370 p.
7. *Muller J.-A.*, *Lemke F.* Self-organizing Data Mining. Extracting Knowledge from Data. – Hamburg: Libri, 2000. – 250 p.
8. *Сарычев А.П.* Идентификация состояний структурно-неопределенных систем. – Днепропетровск: Ин-т техн. механики НАН и НКА Украины, 2008. – 268 с.
9. *Ермаков С.М.*, *Жиглявский А.А.* Математическая теория оптимального эксперимента. – М.: Наука, 1987. – 320 с.
10. *Себер Дж.* Линейный регрессионный анализ. – М.: Мир, 1980. – 456 с.
11. *Сарычев А.П.* Решение проблемы разбиения в МГУА при расчете критерия регулярности в условиях активного эксперимента // Автоматика. – 1989. – № 4. – С. 19–27.
12. *Сарычев А.П.* Определение *J*-оптимального множества регрессоров по повторным выборкам наблюдений // Там же. – 1993. – № 3. – С. 58–66.

Поступила 05.12.2014

Тел. для справок: +38 0562 46-5149 (Днепропетровск)
E-mail: *Sarychev@prognoz.dp.ua*