УДК 519.25:681.5

А.П. Сарычев

УСРЕДНЁННЫЙ КРИТЕРИЙ РЕГУЛЯРНОСТИ В МЕТОДЕ ГРУППОВОГО УЧЁТА АРГУМЕНТОВ

Аннотация. Исследован и аналитически обоснован способ скользящего экзамена, который позволяет определять оптимальное множество регрессоров. Выявлено условие упрощения (редукции) оптимальной регрессионной модели, которое зависит от коэффициентов модели, матрицы наблюдений регрессоров, объемов выборок и дисперсии наблюдений выходной переменной. Оптимальная редуцированная модель будет иметь меньшую ошибку прогнозирования выходной переменной на новых выборках наблюдений по сравнению с истинной моделью.

Ключевые слова: критерий структурной идентификации, условие редукции.

Решение задачи регрессионного анализа в условиях структурной неопределенности по количеству и составу регрессоров предполагает принятие какого-либо способа сравнения моделей, построенных на различных множествах регрессоров. В рамках метода группового учета аргументов (МГУА) [1–4] проведено исследование двух способов сравнения. Первый способ основан на разбиении наблюдений на обучающую и проверочную подвыборки: наблюдения обучающей подвыборки используются для оценивания коэффициентов регрессионной модели, а наблюдения проверочной подвыборки – для оценивания ее качества. Этот способ популярен в практических приложениях, но его применение ограничивает "проблема разбиения" – оптимальное решение может зависеть от разбиения выборки данных на обучающую и проверочную подвыборки. В [4–6] обоснование этого способа проведено в схеме повторных наблюдений, которая возможна в условиях активного эксперимента.

Второй способ – известный способ скользящего экзамена [7–8], в котором в качестве проверочных наблюдений выступают наблюдения, поочередно исключаемые из обучающей выборки – традиционно трактовался как эвристический прием. Аналитическое исследование этого способа, получившего в МГУА название усредненного критерия регулярности, проведено в [9], при этом было установлено, что оптимальное множество регрессоров существует, но зависит от выбора матрицы наблюдений входных переменных. В отличие от работ [4–6], в [9] для усреднённого критерия регулярности не получено условие редукции модели оптимальной сложности. Эта задача является предметом исследования данной статьи.

1. Постановка задачи структурной идентификации в классе статических регрессионных уравнений

Сформулируем задачу структурной идентификации по экспериментальным данным в классе статических регрессионных моделей.

Пусть закон функционирования исследуемого объекта имеет вид

$$y = \overset{o}{y} + \xi = \sum_{j=1}^{\overset{o}{m}} \overset{o}{\theta}_j \overset{o}{x}_j + \xi , \tag{1}$$

где y – выход объекта, измеряемый с ошибкой; $\overset{o}{y}$ – ненаблюдаемый незашумленный выход объекта; ξ – ненаблюдаемая случайная ошибка измерения; $\overset{o}{x}_j$ – j-й вход объекта из множества входов $\overset{o}{X}$, участвующих в формировании выхода объекта ($\overset{o}{X} \neq \varnothing$, $\varnothing$ – пустое множество); $\overset{o}{m}$ – число входов, принадлежащих множеству $\overset{o}{X}$; $\overset{o}{\theta} = (\overset{o}{\theta}_1, \overset{o}{\theta}_2, \ldots, \overset{o}{\theta}_{\overset{o}{m}})^{\mathrm{T}}$ – вектор неизвестных, не равных нулю коэффициентов. Множество входов $\overset{o}{X}$ неизвестно; известно лишь, что $\overset{o}{X} \subseteq X$, где X – некоторое множество точно измеряемых входов объекта, m – их число.

Пусть в результате наблюдения объекта получены: 1) X – $(n \times m)$-матрица n наблюдений m входов множества X, имеющая полный ранг, равный m; 2) y – $(n \times 1)$-вектор, соответствующий наблюдений выхода y. В соответствии с законом (1) для наблюдений должно выполняться

$$\mathrm{y} = \overset{o}{\mathrm{y}} + \xi = \overset{o}{\mathrm{X}} \overset{o}{\theta} + \xi , \tag{2}$$

 ISSN 1562-9945

где $\overset{o}{\mathrm{y}}$ – $(n\times 1)$-вектор значений ненаблюдаемого незашумленного выхода объекта; $\overset{o}{\mathrm{X}}$ – $(n\times m)$-матрица наблюдений входов $\overset{o}{X}$; ξ – $(n\times 1)$-вектор ненаблюдаемых случайных величин.

Пусть относительно ξ выполнены предположения

$$E\{\xi\} = 0_n, \quad E\{\xi\,\xi^{\mathrm{T}}\} = \sigma^2 \cdot \mathrm{I}_n, \tag{3}$$

где $E\{\cdot\}$ – знак математического ожидания по всем возможным реализациям случайного вектора ξ; 0_n – $(n\times 1)$-вектор, состоящий из нулей; σ^2 – неизвестная конечная величина; I_n – единичная $(n\times n)$-матрица.

Требуется найти: 1) множество регрессоров $\overset{o}{X}$, 2) оценку вектора коэффициентов $\overset{o}{\theta}$, 3) оценку дисперсии σ^2.

Для решения поставленной задачи структурного моделирования необходимо: а) указать метод оценивания коэффициентов в моделях с заданной структурой, б) задать алгоритм генерации различных структур моделей, в) принять внешний критерий для оценки качества моделей с разной структурой; г) исследовать поведение математического ожидания внешнего критерия в зависимости от состава регрессоров; д) доказать существование модели оптимальной сложности.

Отметим, что в рассматриваемом классе регрессионных моделей (линейных по входам и коэффициентам) *структура* модели однозначно определяется составом множества входов $V \subseteq X$, присутствующих в модели, а *сложность* модели – их числом.

Пусть $\mathrm{V} - (N\times s)$-матрица наблюдений входов, принадлежащих множеству $V \subseteq X$, s – их число. Условимся называть матрицу X матрицей наблюдений всех регрессоров, матрицу $\overset{o}{\mathrm{X}}$ – матрицей истинного набора регрессоров, а матрицу V – матрицей текущего набора регрессоров, подчеркивая тем самым, что множество входов V меняется в ходе генерации различных структур.

2. Усреднённый критерий регулярности МГУА

Усредненный критерий регулярности в методе группового учета аргументов (МГУА) [2, 4, 9] применяется в тех случаях, когда по каким-либо причинам невозможно разбиение исходной выборки наблюдений на обучающую и проверочную подвыборки.

Пусть поставлена задача структурного моделирования (1)–(3). Рассмотрим схему решения этой задачи, состоящую в полном переборе всевозможных структур моделей. На каждом этапе алгоритма с номером s $(s = 1, 2, \ldots, m)$ рассматриваются все возможные модели сложности s

$$\hat{y} = \sum_{j \in J^s} \hat{d}_j x_j, \tag{4}$$

где J^s – одно из всевозможных подмножеств множества $\{1, 2, \ldots, m\}$ такое, что число элементов в нем равно s; число таких всевозможных подмножеств равно C_m^s – числу сочетаний из m элементов по s.

Пусть $(s \times 1)$-вектор $\hat{\mathrm{d}}$ – полученная по методу наименьших квадратов оценка вектора коэффициентов регрессии y по V

$$\mathrm{y} = \mathrm{Vd} + \mathrm{e}, \tag{5}$$

рассчитанная по всем имеющимся n наблюдениям:

$$\hat{\mathrm{d}} = (\mathrm{V}^{\mathrm{T}}\mathrm{V})^{-1}\mathrm{V}^{\mathrm{T}}\mathrm{y} = \arg\min_{\mathrm{d}} \Phi(\mathrm{d}), \tag{6}$$

$$\Phi(\mathrm{d}) = \mathrm{e}^{\mathrm{T}}\mathrm{e} = (\mathrm{y} - \mathrm{Vd})^{\mathrm{T}}(\mathrm{y} - \mathrm{Vd}). \tag{7}$$

Пусть $\hat{\mathrm{d}}_{(i)}$ – МНК-оценка регрессии, рассчитанная по выборке, которая получается из исходной выборки в результате исключения из нее наблюдения i:

$$\hat{\mathrm{d}}_{(i)} = (\mathrm{V}_{(i)}^{\mathrm{T}}\mathrm{V}_{(i)})^{-1}\mathrm{V}_{(i)}^{\mathrm{T}}\mathrm{y}_{(i)}. \tag{8}$$

где для $((n-1) \times s)$-матрицы $\mathrm{V}_{(i)}$ и для $((n-1) \times 1)$-вектора $\mathrm{y}_{(i)}$ выполняется

$$\mathrm{V}^{\mathrm{T}} = [\mathrm{V}_{(i)}^{\mathrm{T}} \mid \mathrm{v}_i], \quad \mathrm{y}^{\mathrm{T}} = (\mathrm{y}_{(i)}^{\mathrm{T}}, y_i), \tag{9}$$

а v_i – $(s \times 1)$-вектор значений входов, соответствующих y_i – наблюдению выхода с номером i (выполнение (9) для любого i можно обеспечить простой перестановкой столбцов матрицы V^{T} и элементов вектора y).

Случайная величина

$$УКР(V) = \sum_{i=1}^{n} (y_i - \mathrm{v}_i^{\mathrm{T}} \hat{\mathrm{d}}_{(i)})^2 \tag{10}$$

называется в МГУА усредненным критерием регулярности.

Учитывая (6)–(9) и применяя правило изменения обратной матрицы при малоранговой модификации [10, с. 31–32] (формула

Бартлетта), получаем так называемую формулу рекуррентного оценивания по МНК:

$$\hat{d} = \hat{d}_{(i)} + (1 + v_i^T P_{(i)} v_i)^{-1} P_{(i)} v_{(i)} (y_i - v_i^T \hat{d}_{(i)}), \tag{11}$$

$$P_{(i)} = [V_{(i)}^T V_{(i)}]^{-1} = [V^T V - v_i v_i^T]^{-1}. \tag{12}$$

Далее, после несложных преобразований (11), получаем оценку $\hat{d}_{(i)}$, выраженную через величины $\hat{d}$, $P_{(i)}$, v_i и y_i:

$$\hat{d}_{(i)} = \hat{d} + \left[P_{(i)} v_i v_i^T - I_s \cdot (1 + v_i^T P_{(i)} v_i) \right]^{-1} P_{(i)} v_i (y_i - v_i^T \hat{d}) =$$

$$= \hat{d} + K_{(i)}^{-1} P_{(i)} v_i (y_i - v_i^T \hat{d}), \tag{13}$$

$$K_{(i)} = \left[P_{(i)} v_i v_i^T - I_s \cdot (1 + v_i^T P_{(i)} v_i) \right]^{-1}. \tag{14}$$

Учитывая (13)–(14), для (10) получаем

$$УКР(V) = \sum_{i=1}^{n} \left\{ y_i - v_i^T [\hat{d} + K_{(i)}^{-1} P_{(i)} v_i (y_i - v_i^T \hat{d})] \right\}^2 =$$

$$= \sum_{i=1}^{n} \left\{ (1 - v_i^T K_{(i)}^{-1} P_{(i)} v_i)^2 (y_i - v_i^T \hat{d}) \right\}^2 = \sum_{i=1}^{n} \omega_i^{-2}(V) \cdot (y_i - v_i^T \hat{d})^2 =$$

$$= (y - V\hat{d})^T \Omega(V)(y - V\hat{d}) = (y - \hat{y})^T \Omega(V)(y - \hat{y}), \tag{15}$$

$$\Omega(V) = \{ \omega_1^{-2}(V), \omega_2^{-2}(V), \ldots, \omega_n^{-2}(V) \}, \tag{16}$$

где $\hat{y} = V\hat{d}$ – выход регрессионной модели; $\omega_i^{-2}(V)$, $i = 1, 2, \ldots, n$ – элементы, для которых выполняется

$$\omega_i^{-2}(V) = (1 - v_i^T K_{(i)}^{-1} P_{(i)} v_i)^2 = (1 + v_i^T P_{(i)} v_i)^2 = (1 - v_i^T (V^T V)^{-1} v_i)^{-2}, \tag{17}$$

т. е. они определяются *исключительно* наблюдениями регрессоров; если $v_i \neq 0_s$, то выполняется $\omega_i^{-2}(V) > 1$, а если $v_i = 0_s$, то $\omega_i^{-2}(V) = 1$.

Учитывая (6), (15)–(17) и введя симметричную идемпотентную матрицу $S_V = I_n - V(V^T V)^{-1} V^T$, для усредненного критерия регулярности получаем

$$УКР(V) = \overset{o}{y}{}^T S_V \Omega_V S_V \overset{o}{y} + 2\overset{o}{y}{}^T S_V \Omega_V S_V \xi + \xi^T S_V \Omega_V S_V \xi. \tag{18}$$

Учитывая предположения (2)–(3), для математического ожидания и дисперсии усредненного критерия регулярности получаем

$$E\{УКР(V)\} = \overset{o}{y}{}^T S_V \Omega_V S_V \overset{o}{y} + \sigma^2 \cdot \mathrm{tr}\, [\Omega_V^{1/2}], \tag{19}$$

$$D\{УКР(V)\} = 4\sigma^2 \overset{o}{\mathrm{y}}^{\mathrm{T}} \mathrm{S}_V \Omega_V \mathrm{S}_V \Omega_V \mathrm{S}_V \overset{o}{\mathrm{y}} + 2\sigma^4 \cdot \mathrm{tr}\,[\,\mathrm{S}_V \Omega_V \mathrm{S}_V \Omega_V], \qquad (20)$$

где при расчете дисперсии дополнительно предположено, что случайные величины ξ_i, $i = 1, 2, \ldots, n$, имеют нормальное распределение: $\xi \sim N_n(0_n, \sigma^2 \mathrm{I}_n)$.

В случае истинной структуры для математического ожидания усредненного критерия регулярности получаем

$$E\{УКР(\overset{o}{X})\} = \sigma^2 \cdot \mathrm{tr}\,[\,\Omega^{1/2}_{\overset{o}{X}}]\,. \qquad (21)$$

В [9] проведено аналитическое исследование усредненного критерия регулярности и показано, что он позволяет находить регрессионную модель оптимальной сложности. Для анализа *УКР* (с точки зрения возможности применения этой статистики для решения задачи структурного идентификации) в схеме полного перебора всех возможных структур моделей в [9] введены множества входов R, $\overline{X}$, M ($\varnothing$ – пустое множество)

$$V = R \cup \overline{X}, \quad \overset{o}{X} = M \cup \overline{X}, \quad R \cap \overline{X} = \varnothing, \quad R \cap M = \varnothing, \qquad (22)$$

и соответствующие матрицы наблюдений

$$\mathrm{V} = [\mathrm{R} \mid \overline{\mathrm{X}}], \quad \overset{o}{\mathrm{X}} = [\mathrm{M} \mid \overline{\mathrm{X}}], \qquad (23)$$

где R – $(n \times l)$-матрица избыточных регрессоров множества входов R; M – $(n \times p)$-матрица недостающих (пропущенных) регрессоров множества входов M; $\overline{\mathrm{X}}$ – $(n \times \overline{m})$-матрица истинных регрессоров таких, что для множества $\overline{X}$ выполнено $\overline{X} \subseteq \overset{o}{X}$; $l, p, \overline{m}$ – целые положительные числа такие, что $s = l + \overline{m}$, $\overset{o}{m} = p + \overline{m}$. Таким образом, текущее множество входов V отличается от множества истинных входов $\overset{o}{X}$ тем, что в последнем часть истинных входов (множество M) замещена множеством избыточных входов (множество R), не участвующих в формировании выхода объекта.

В [9] установлено: а) первый член в (19) обладает свойствами, требуемыми от статистики, предназначенной для поиска оптимального множества регрессоров – он положителен, если в анализируемом текущем множестве регрессоров недостает части истинных регрессоров ($M \neq \varnothing$), и он равен нулю, если недостающих регрессоров нет ($M = \varnothing$); б) второй член в (19) обладает свойствами функции штрафа

 ISSN 1562-9945

за “переусложнение” структуры модели – он возрастает при добавлении новых регрессоров.

3. Условие редукции модели оптимальной сложности для усреднённого критерия регулярности

Продолжим исследование свойств усредненного критерия регулярности МГУА, проведённое в [9]. Используя (4)–(23), получим условие редукции – условие упрощения регрессионной модели, оптимальной по усредненному критерию регулярности.

Случай недостающего регрессора. Пусть в модель ошибочно не включен один регрессор, хотя он участвует в формировании значения выходной переменной, и для простоты будем считать, что это регрессор с номером $\overset{o}{m}$ из множества $\overset{o}{X}$. Тогда для истинного и текущего множества регрессоров, их матриц наблюдений и наблюдений с номером i, выполняется

$$\overset{o}{X} = V \cup \overset{o}{x}(\overset{o}{m}), \qquad \overset{o}{\mathrm{X}} = [\ \mathrm{V} \mid \mathrm{m}\ ], \qquad \overset{o}{\mathrm{X}}{}^{\mathrm{T}} = \begin{bmatrix} \mathrm{V}^{\mathrm{T}} \\ \hline \mathrm{m}^{\mathrm{T}} \end{bmatrix}, \qquad \overset{o}{\mathrm{x}}{}_i^{\mathrm{T}} = \left(\mathrm{v}_i^{\mathrm{T}} \mid m_i \right),$$

$$\overset{o}{\mathrm{x}}_i = \begin{pmatrix} \mathrm{v}_i \\ \hline m_i \end{pmatrix}, \tag{24}$$

где $\overset{o}{x}(\overset{o}{m})$ – пропущенный вход, m – соответствующий ему регрессор, m_i – i-е наблюдение регрессора m.

Введем обозначение

$$\overset{o}{\theta} = \begin{pmatrix} \overset{o}{\theta}(V) \\ \hline \overset{o}{\theta}(\overset{o}{m}) \end{pmatrix}. \tag{25}$$

Учитывая результаты (19) и (21), запишем разность

$$\Delta_1(V, \overset{o}{X}) = E\{УКР(V)\} - E\{УКР(\overset{o}{X})\} =$$

$$= \overset{o}{\mathrm{y}}{}^{\mathrm{T}} \mathrm{S}_V \Omega_V \mathrm{S}_V \overset{o}{\mathrm{y}} + \sigma^2 \cdot \mathrm{tr}\,[\,\Omega_V^{1/2}] - \sigma^2 \cdot \mathrm{tr}\,[\,\Omega_{\overset{o}{X}}^{1/2}]. \tag{26}$$

Для первой составляющей в (26), обусловленной ошибкой в выборе структуры (24), выполняется

$$\overset{o}{\mathrm{y}}{}^{\mathrm{T}} \mathrm{S}_V \Omega_V \mathrm{S}_V \overset{o}{\mathrm{y}} =$$

$$= \overset{o}{\theta}{}^{\mathrm{T}} \begin{bmatrix} \mathrm{V}^{\mathrm{T}} \\ \hline \mathrm{m}^{\mathrm{T}} \end{bmatrix} [\mathrm{I}_n - \mathrm{V}(\mathrm{V}^{\mathrm{T}}\mathrm{V})^{-1}\mathrm{V}^{\mathrm{T}}]\, \Omega_V\, [\mathrm{I}_n - \mathrm{V}(\mathrm{V}^{\mathrm{T}}\mathrm{V})^{-1}\mathrm{V}^{\mathrm{T}}]\, [\ \mathrm{V} \mid \mathrm{m}\ ]\, \overset{o}{\theta} =$$

$$= \overset{\circ}{\theta}{}^{T}\left[\frac{O_{s\times n}}{m^{T}[I_n - V(V^{T}V)^{-1}V^{T}]}\right]\Omega_V\left[\,[O_{n\times s}]\;\middle|\;[I_n - V(V^{T}V)^{-1}V^{T}]\,m\,\right]\overset{\circ}{\theta} =$$

$$= \left(\overset{\circ}{\theta}{}^{T}(V), \overset{\circ}{\theta}(\overset{\circ}{m})\right)\left[\begin{array}{c|c} O_{s\times s} & 0_s \\ \hline 0_s^{T} & \begin{array}{c} m^{T}[I_n - V(V^{T}V)^{-1}V^{T}]\times \\ \times\Omega_V[I_n - V(V^{T}V)^{-1}V^{T}]m \end{array}\end{array}\right]\left(\begin{array}{c}\overset{\circ}{\theta}(V) \\ \hline \overset{\circ}{\theta}(\overset{\circ}{m})\end{array}\right) =$$

$$= \overset{\circ}{\theta}(\overset{\circ}{m})m^{T}[I_n - V(V^{T}V)^{-1}V^{T}]\,\Omega_V[I_n - V(V^{T}V)^{-1}V^{T}]\,m\,\overset{\circ}{\theta}(\overset{\circ}{m}) =$$

$$= \overset{\circ}{\theta}(\overset{\circ}{m})m^{T}S_V\,\Omega_V\,S_V\,m\,\overset{\circ}{\theta}(\overset{\circ}{m}). \tag{27}$$

Учитывая (24), для разности следов в (26) получаем

$$a_1 = \sigma^2\cdot \mathrm{tr}\,[\,\Omega_V^{1/2}] - \sigma^2\cdot \mathrm{tr}\,[\,\Omega_{\overset{\circ}{X}}^{1/2}] =$$

$$= \sigma^2\cdot\sum_{i=1}^{n}\left(1 - v_i^{T}(V^{T}V)^{-1}v_i\right)^{-1} - \sigma^2\cdot\sum_{i=1}^{n}\left\{1 - \left(v_i^{T}\;\middle|\;m_i\right)\left[\left[\frac{V^{T}}{m^{T}}\right][\,V\;|\;m\,]\right]^{-1}\left(\frac{v_i}{m_i}\right)\right\}^{-1} =$$

$$= \sigma^2\cdot\sum_{i=1}^{n}\left\{1 - v_i^{T}(V^{T}V)^{-1}v_i\right\}^{-1} - \sigma^2\cdot\sum_{i=1}^{n}\left\{1 - \left(v_i^{T}\;\middle|\;m_i\right)\left[\begin{array}{c|c} V^{T}V & V^{T}m \\ \hline m^{T}V & m^{T}m\end{array}\right]^{-1}\left(\frac{v_i}{m_i}\right)\right\}^{-1}$$

. (28)

Для перемножения блочных матриц в (28) применим формулу обращения блочной матрицы (частный случай формулы Фробениуса [11, с. 302]):

$$A^{-1} = \left[\begin{array}{c|c} B & c \\ \hline d^{T} & e\end{array}\right]^{-1} = \left[\begin{array}{c|c} B^{-1} + B^{-1}c\cdot f^{-1}\cdot d^{T}B^{-1} & -B^{-1}c\cdot f^{-1} \\ \hline -f^{-1}\cdot d^{T}B^{-1} & f^{-1}\end{array}\right],$$

$$f = e - d^{T}B^{-1}c \tag{29}$$

(в нашем случае выполняется $B = V^{T}V$, $c = V^{T}m$, $d^{T} = m^{T}V$, $e = m^{T}m$).

Учитывая (29), для (28) получаем

$$a_1 = \sigma^2\cdot\sum_{i=1}^{n}(1 - v_i^{T}(V^{T}V)^{-1}v_i)^{-1} - \sigma^2\cdot\sum_{i=1}^{n}(1 - v_i^{T}(V^{T}V)^{-1}v_i - g_i^2)^{-1}, \tag{30}$$

$$g_i^2 = f^{-1}\cdot(m_i - v_i^{T}(V^{T}V)^{-1}V^{T}m)\cdot(m_i - v_i^{T}(V^{T}V)^{-1}V^{T}m), \tag{31}$$

$$f = m^{T}m - m^{T}V(V^{T}V)^{-1}V^{T}m = m^{T}(I_n - V(V^{T}V)^{-1}V^{T})\,m. \tag{32}$$

 ISSN 1562-9945

Применяя формулу разложения функции $\frac{1}{1-x}$ в ряд при малых x, пренебрегая членами второго и более высокого порядков, получаем

$$a_1 = \sigma^2 \cdot \mathrm{tr}\,[\,\Omega_V^{1/2}] - \sigma^2 \cdot \mathrm{tr}\,[\,\Omega_{\overset{o}{X}}^{1/2}] \approx$$

$$\approx \sigma^2 \cdot \sum_{i=1}^{n} \left(1 + \mathrm{v}_i^{\mathrm{T}}(\mathrm{V}^{\mathrm{T}}\mathrm{V})^{-1}\mathrm{v}_i\right) - \sigma^2 \cdot \sum_{i=1}^{n} \left(1 + \mathrm{v}_i^{\mathrm{T}}(\mathrm{V}^{\mathrm{T}}\mathrm{V})^{-1}\mathrm{v}_i\right) -$$

$$-\sigma^2 \cdot f^{-1} \cdot \sum_{i=1}^{n} (m_i - \mathrm{v}_i^{\mathrm{T}}(\mathrm{V}^{\mathrm{T}}\mathrm{V})^{-1}\mathrm{V}^{\mathrm{T}}\mathrm{m}) \cdot (m_i - \mathrm{v}_i^{\mathrm{T}}(\mathrm{V}^{\mathrm{T}}\mathrm{V})^{-1}\mathrm{V}^{\mathrm{T}}\mathrm{m}) =$$

$$= \sigma^2 \cdot (n + \overset{o}{m} - 1) - \sigma^2 \cdot (n + \overset{o}{m} - 1) - \sigma^2 = -\sigma^2. \qquad (33)$$

Объединяя (27) и (33), получаем

$$\Delta_1(V, \overset{o}{X}) = E\{УКР(V)\} - E\{УКР(\overset{o}{X})\}$$

$$= \overset{o}{\theta}(\overset{o}{m})\mathrm{m}^{\mathrm{T}}\mathrm{S}_V\,\Omega_V\,\mathrm{S}_V\,\mathrm{m}\,\overset{o}{\theta}(\overset{o}{m}) - \sigma^2. \qquad (34)$$

Выполнение $\Delta_1(V, \overset{o}{X}) \le 0$ является условием так называемой редукции модели, оптимальной по структуре. Из (34) для условия редукции получаем

$$\overset{o}{\theta}(\overset{o}{m})\mathrm{m}^{\mathrm{T}}\mathrm{S}_V\,\Omega_V\,\mathrm{S}_V\,\mathrm{m}\,\overset{o}{\theta}(\overset{o}{m}) \le \sigma^2. \qquad (35)$$

Отличие (35) от результатов [4–6] состоит в присутствии диагональной весовой матрицы Ω_V.

Редукция модели, оптимальной по составу регрессоров, означает, что при выполнении соотношения между параметрами модели (35) следует исключить регрессор m из модели. Редуцированная модель будет иметь меньшую ошибку прогнозирования выходной переменной на новых выборках наблюдений по сравнению с истинной моделью.

Из (35) следует, что возможность редукции модели может быть обусловлена пятью причинами: а) малой нормой коэффициента $\overset{o}{\theta}(\overset{o}{m})$; б) малой нормой вектора наблюдений регрессора m; в) малым объёмом выборок наблюдений n; г) высокой степенью линейной зависимости регрессора m с другими регрессорами в матрице V; д) большим значением дисперсии σ^2.

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

$$V = \overset{\circ}{X} \cup r, \qquad \mathrm{V} = \left[\overset{\circ}{\mathrm{X}} \,\vdots\, \mathrm{r} \right], \qquad \mathrm{V}^{\mathrm{T}} = \begin{bmatrix} \overset{\circ}{\mathrm{X}}{}^{\mathrm{T}} \\ \hline \mathrm{r}^{\mathrm{T}} \end{bmatrix}, \qquad \mathrm{v}_i^{\mathrm{T}} = \left(\overset{\circ}{\mathrm{x}}{}_i^{\mathrm{T}} \,\vdots\, r_i \right),$$

$$\mathrm{v}_i = \begin{pmatrix} \overset{\circ}{\mathrm{x}}_i \\ \hline r_i \end{pmatrix}, \tag{36}$$

где r – излишний вход, а r – соответствующий ему избыточный регрессор.

В этом случае введённая в (18) составляющая, обусловленная выбором текущего множества регрессоров V вместо истинного множества $\overset{\circ}{X}$, равна нулю. Действительно, учитывая (36), получаем

$$\overset{\circ}{\theta}{}^{\mathrm{T}} \overset{\circ}{\mathrm{X}}{}^{\mathrm{T}} \mathrm{S}_V \Omega_V \mathrm{S}_V \overset{\circ}{\mathrm{X}} \overset{\circ}{\theta} = \overset{\circ}{\theta}{}^{\mathrm{T}} \overset{\circ}{\mathrm{X}}{}^{\mathrm{T}} \left[\mathrm{I}_n - \mathrm{V}(\mathrm{V}^{\mathrm{T}}\mathrm{V})^{-1}\mathrm{V}^{\mathrm{T}} \right] \Omega_V \mathrm{S}_V \overset{\circ}{\mathrm{X}} \overset{\circ}{\theta} =$$

$$= \overset{\circ}{\theta}{}^{\mathrm{T}} \overset{\circ}{\mathrm{X}}{}^{\mathrm{T}} \left[\mathrm{I}_n - \left[\overset{\circ}{\mathrm{X}} \,\vdots\, \mathrm{r} \right] \left(\begin{bmatrix} \overset{\circ}{\mathrm{X}}{}^{\mathrm{T}}\overset{\circ}{\mathrm{X}} & \overset{\circ}{\mathrm{X}}{}^{\mathrm{T}}\mathrm{r} \\ \mathrm{r}^{\mathrm{T}}\overset{\circ}{\mathrm{X}} & \mathrm{r}^{\mathrm{T}}\mathrm{r} \end{bmatrix} \right)^{-1} \begin{bmatrix} \overset{\circ}{\mathrm{X}}{}^{\mathrm{T}} \\ \hline \mathrm{r}^{\mathrm{T}} \end{bmatrix} \right] \Omega_V \mathrm{S}_V \overset{\circ}{\mathrm{X}} \overset{\circ}{\theta} =$$

$$= \overset{\circ}{\theta}{}^{\mathrm{T}} \mathrm{O}_{(s\times s)} \Omega_V \mathrm{S}_V \overset{\circ}{\mathrm{X}} \overset{\circ}{\theta} = 0, \tag{37}$$

где при перемножении блочных матриц применена формула обращения блочной матрицы (частный случай формулы Фробениуса [11, с. 302]).

Тогда, для случая избыточного регрессора, учитывая (36)–(37) и проводя вычисления аналогично (28)–(33), получаем

$$\Delta_2(V, \overset{\circ}{X}) = E\{УКР(V)\} - E\{УКР(\overset{\circ}{X})\} = \sigma^2 \cdot \mathrm{tr}\,[\,\Omega_V^{1/2}] - \sigma^2 \cdot \mathrm{tr}\,[\,\Omega_{\overset{\circ}{X}}^{1/2}] =$$

$$= \sigma^2 \cdot \sum_{i=1}^{n} \left(1 - \mathrm{v}_i^{\mathrm{T}}(\mathrm{V}^{\mathrm{T}}\mathrm{V})^{-1}\mathrm{v}_i \right)^{-1} - \sigma^2 \cdot \sum_{i=1}^{n} \left(1 - \overset{\circ}{\mathrm{x}}{}_i^{\mathrm{T}}(\overset{\circ}{\mathrm{X}}{}^{\mathrm{T}}\overset{\circ}{\mathrm{X}})^{-1}\overset{\circ}{\mathrm{x}}_i \right)^{-1} =$$

$$= \sigma^2 \cdot \sum_{i=1}^{n} \left\{ 1 - \left(\overset{\circ}{\mathrm{x}}{}_i^{\mathrm{T}} \,\vdots\, r_i \right) \begin{bmatrix} \overset{\circ}{\mathrm{X}}{}^{\mathrm{T}}\overset{\circ}{\mathrm{X}} & \overset{\circ}{\mathrm{X}}{}^{\mathrm{T}}\mathrm{r} \\ \mathrm{r}^{\mathrm{T}}\overset{\circ}{\mathrm{X}} & \mathrm{r}^{\mathrm{T}}\mathrm{r} \end{bmatrix}^{-1} \begin{pmatrix} \overset{\circ}{\mathrm{x}}_i \\ \hline r_i \end{pmatrix} \right\}^{-1} - \sigma^2 \cdot \sum_{i=1}^{n} \left(1 - \overset{\circ}{\mathrm{x}}{}_i^{\mathrm{T}}(\overset{\circ}{\mathrm{X}}{}^{\mathrm{T}}\overset{\circ}{\mathrm{X}})^{-1}\overset{\circ}{\mathrm{x}}_i \right)^{-1} =$$

$$= \sigma^2 \cdot \sum_{i=1}^{n} (1 - \overset{\circ}{\mathrm{x}}{}_i^{\mathrm{T}}(\overset{\circ}{\mathrm{X}}{}^{\mathrm{T}}\overset{\circ}{\mathrm{X}})^{-1}\overset{\circ}{\mathrm{x}}_i - q_i^2)^{-1} - \sigma^2 \cdot \sum_{i=1}^{n} (1 - \overset{\circ}{\mathrm{x}}{}_i^{\mathrm{T}}(\overset{\circ}{\mathrm{X}}{}^{\mathrm{T}}\overset{\circ}{\mathrm{X}})^{-1}\overset{\circ}{\mathrm{x}}_i)^{-1}, \tag{38}$$

 ISSN 1562-9945

где

$$q_i^2 = f^{-1} \cdot (r_i - \overset{\circ}{x}{}_i^{\mathrm{T}}(\overset{\circ}{X}{}^{\mathrm{T}}\overset{\circ}{X})^{-1}\overset{\circ}{X}{}^{\mathrm{T}}\mathrm{r}) \cdot (r_i - \overset{\circ}{x}{}_i^{\mathrm{T}}(\overset{\circ}{X}{}^{\mathrm{T}}\overset{\circ}{X})^{-1}\overset{\circ}{X}{}^{\mathrm{T}}\mathrm{r}), \tag{39}$$

$$f = \mathrm{r}^{\mathrm{T}}\mathrm{r} - \mathrm{r}^{\mathrm{T}}\overset{\circ}{X}(\overset{\circ}{X}{}^{\mathrm{T}}\overset{\circ}{X})^{-1}\overset{\circ}{X}{}^{\mathrm{T}}\mathrm{r} = \mathrm{r}^{\mathrm{T}}(\mathrm{I}_n - \overset{\circ}{X}(\overset{\circ}{X}{}^{\mathrm{T}}\overset{\circ}{X})^{-1}\overset{\circ}{X}{}^{\mathrm{T}})\,\mathrm{r}. \tag{40}$$

Применяя формулу разложения функции $\frac{1}{1-x}$ в ряд при малых x, пренебрегая членами второго и более высокого порядков, получаем

$$\Delta_2(V, \overset{\circ}{X}) \approx \sigma^2 \cdot \sum_{i=1}^{n} (1 + \overset{\circ}{x}{}_i^{\mathrm{T}}(\overset{\circ}{X}{}^{\mathrm{T}}\overset{\circ}{X})^{-1}\overset{\circ}{x}{}_i) +$$

$$+\sigma^2 \cdot f^{-1} \cdot \sum_{i=1}^{n} (r_i - \overset{\circ}{x}{}_i^{\mathrm{T}}(\overset{\circ}{X}{}^{\mathrm{T}}\overset{\circ}{X})^{-1}\overset{\circ}{X}{}^{\mathrm{T}}\mathrm{r}) \cdot (r_i - \overset{\circ}{x}{}_i^{\mathrm{T}}(\overset{\circ}{X}{}^{\mathrm{T}}\overset{\circ}{X})^{-1}\overset{\circ}{X}{}^{\mathrm{T}}\mathrm{r}) -$$

$$-\sigma^2 \cdot \sum_{i=1}^{n} (1 + \overset{\circ}{x}{}_i^{\mathrm{T}}(\overset{\circ}{X}{}^{\mathrm{T}}\overset{\circ}{X})^{-1}\overset{\circ}{x}{}_i) =$$

$$= \sigma^2 \cdot (n + \overset{\circ}{m}) + \sigma^2 - \sigma^2 \cdot (n + \overset{\circ}{m}) = \sigma^2. \tag{41}$$

Из (41) следует, что в случае избыточного регрессора истинная структура всегда лучше, а регрессор r действительно не следует включать в модель.

Выводы

В статье исследован и аналитически обоснован способ скользящего экзамена для сравнения регрессионных моделей, построенных на различных множествах регрессоров. Несмотря на успешное применение этого способа на практике и неоднократное подтверждение его работоспособности методом статистических испытаний, он традиционно считается эвристическим приёмом.

Получены условия существования оптимального множества регрессоров, зависящие от коэффициентов моделей, матриц наблюдений регрессоров и объемов выборок, и выявлены закономерности упрощения (редукции) оптимальной регрессионной модели при уменьшении объемов выборок и при увеличении дисперсии наблюдений. При выполнении условия редукции редуцированная модель будет иметь меньшую ошибку прогнозирования выходной переменной на новых выборках наблюдений по сравнению с истинной моделью.

ЛИТЕРАТУРА

1. Ивахненко А. Г. Индуктивный метод самоорганизации моделей сложных систем / А. Г. Ивахненко. – К. : Наук. думка, 1982. – 296 с.
2. Ивахненко А. Г. Помехоустойчивость моделирования / А. Г. Ивахненко, В. С. Степашко. – Киев : Наукова думка, 1985. – 216 с.
3. Ивахненко А. Г. Моделирование сложных систем по экспериментальным данным / А. Г. Ивахненко А. Г., Ю. П. Юрачковский. – М.: Радио и связь, 1987. – 120 с.
4. Сарычев А. П. Идентификация состояний структурно-неопределенных систем / А. П. Сарычев. – Днепропетровск : НАН Украины и НКА Украины, Институт технической механики, 2008. – 268 с.
5. Сарычев А. П. Решение проблемы разбиения в МГУА при расчете критерия регулярности в условиях активного эксперимента / А. П. Сарычев // Автоматика. – 1989. – № 4. – С. 19–27.
6. Сарычев А. П. Определение J-оптимального множества регрессоров по повторным выборкам наблюдений / А. П. Сарычев // Автоматика. – 1993. – № 3. – С. 58–66.
7. Hocking R. R. The Analysis and Selection of Variables in Linear Regression / R. R. Hocking // Biometrics. – 1976. – Vol. 32. – P. 1–49.
8. Вапник В. Н. Восстановление зависимостей по эмпирическим данным / В. Н. Вапник. – М. : Наука, 1979. – 448 с.
9. Сарычев А. П. Усредненный критерий регулярности метода группового учета аргументов в задаче поиска наилучшей регрессии / А. П. Сарычев // Автоматика. – 1990. – № 5. – С. 28–33.
10. Хорн Р. Матричный анализ / Р. Хорн, Ч. Джонсон; пер. с англ. – М.: Мир, 1989. – 656 с.
11. Ермаков С. М. Математическая теория оптимального эксперимента / С. М. Ермаков , А. А. Жиглявский. – М.: Наука, 1987. – 320с.

 ISSN 1562-9945