

*М. П. Базилевский*<sup>1</sup>

*1 Иркутский государственный университет путей сообщения, г. Иркутск, Российская Федерация*

## **ФУНДАМЕНТАЛЬНЫЙ БЛОК АЛГОРИТМОВ ПОСТРОЕНИЯ ХОРОШО ИНТЕРПРЕТИРУЕМЫХ КАЧЕСТВЕННЫХ РЕГРЕССИОННЫХ МОДЕЛЕЙ**

**Аннотация.** Статья посвящена разработке алгоритма построения качественных и хорошо интерпретируемых регрессионных моделей. Проведен анализ программного комплекса автоматизации процесса построения регрессионных моделей, предназначенного для реализации "конкурса" моделей. На основе анализа выявлен ряд недостатков данного программного продукта, в частности, то, что полученные в результате его работы регрессионные модели зачастую трудно интерпретировать. На основе выявленных недостатков впервые введено понятие "хорошо интерпретируемая качественная модель". Сформулированы требования к таким регрессиям. Рассмотрен эффективный алгоритм реализации "конкурса" моделей "Selection B". Разработан фундаментальный блок алгоритмов построения хорошо интерпретируемых качественных регрессионных моделей. Для его реализации достаточно модернизировать эффективный алгоритм "Selection B".

**Ключевые слова:** регрессионный анализ, хорошо интерпретируемая качественная регрессионная модель, "конкурс" моделей, интерпретация, мультиколлинеарность, алгоритм.

*М.Р. Bazilevskiy*<sup>1</sup>

*1 Irkutsk State Transport University, Irkutsk, Russia*

## **THE FUNDAMENTAL BLOCK OF ALGORITHMS FOR CONSTRUCTING WELL-INTERPRETED QUALITATIVE REGRESSION MODELS**

**Abstract.** The article is devoted to the development of an algorithm for constructing high-quality and well-interpreted regression models. The analysis of the program complex for the automation of the process of constructing regression models, intended for the implementation of the "competition" of models, is performed. Based on the analysis, a number of shortcomings of this software product were identified, in particular, the fact that the regression models obtained as a result of its work are often difficult to interpret. On the basis of the identified shortcomings, the concept of "well-interpreted qualitative model" was introduced for the first time. Requirements for such regressions are formulated. An efficient algorithm "Selection B" for the implementation of the "competition" models is considered. A fundamental block of algorithms for constructing well-interpreted qualitative regression models has been developed. To implement it, it is enough to modernize the effective algorithm "Selection B".

**Keywords:** regression analysis, well-interpreted qualitative regression model, "competition" of models, interpretation, multicollinearity, algorithm.

**Введение.** Качественным способом, позволяющим справляться с проблемой выбора спецификации регрессии, является организация "конкурса" моделей, состоящего в формировании множества их альтернативных вариантов и последующем выборе наиболее приемлемого из них на основе одного или нескольких критериев адекватности. Алгоритмическая схема организации "конкурса" моделей разработана профессором С.И. Носковым и подробно описана в его монографии [1]. Для автоматизации этой схемы было разработано специализированное программное обеспечение – программный комплекс автоматизации процесса построения регрессионных моделей (ПК АППРМ) [2, 3]. С помощью ПК АППРМ успешно решено множество прикладных задач анализа данных. Так, в работе [3] проводилось моделирование работы выпарного аппарата на большом промышленном предприятии, в работах [4, 5] – валового регионального продукта Иркутской области, в [6–8] – динамики эксплуатационных показателей функционирования Красноярской железной дороги, в [9] – обстановки с пожарами в сельских населенных пунктах в условиях их газификации, в [10] – уровня безработицы в России и т.д.

Проведенный автором анализ ПК АППРМ позволил выявить его некоторые недостатки.

1. Полученные с помощью ПК АППРМ аддитивные и линейно-мультипликативные модели прекрасно подходят для прогнозирования, однако в реальности представляют собой сложные нелинейные уравнения, коэффициенты которых весьма затруднительно интерпретировать.

2. ПК АППРМ, разработанный в среде программирования Borland C++ Builder, довольно долго осуществляет процедуру выбора наилучшей зависимости. На сегодняшний день существуют более эффективные программные средства для проведения регрессионного анализа, например, эконометрический пакет Gretl, имеющий встроенный скриптовый язык hansl.

3. В ПК АППРМ полностью игнорируется проблема мультиколлинеарности, которая может лишать смысла интерпретацию коэффициентов модели.

4. В ПК АППРМ из множества альтернатив выбирается только одна единственная регрессия. При этом за гранью исследования остаются миллионы других регрессий, многие из которых и вовсе незначительно отличаются от наилучшей модели. Поэтому при организации «конкурса» моделей необходимо подкреплять его результаты статистическим анализом критериальной матрицы [11].

5. В ПК АППРМ не проверяется соответствие знаков регрессионных коэффициентов физическому смыслу входящих в уравнение факторов. В монографии [1] отмечено, что все варианты уравнений, какими бы высокими значениями критериев они не характеризовались, необходимо исключить из участия в «конкурсе» в случае нарушения этого соответствия хотя бы по одному параметру.

6. В ПК АППРМ не реализованы алгоритмы, исключающие незначимые по величине  $t$ -критерия Стьюдента коэффициенты регрессий.

На основе выявленных недостатков впервые введем понятие "хорошо интерпретируемая качественная регрессионная модель" или ХИК-модель.

*Определение.* Регрессионная модель называется хорошо интерпретируемой и качественной (ХИК), если она удовлетворяет следующим пяти условиям.

1. Её спецификация изначально выбрана так, что после оценивания модели можно интерпретировать любой её коэффициент или некоторый его аналог, за исключением, быть может, свободного члена.

2. Знаки коэффициентов оцененной модели соответствуют физическому смыслу входящих в уравнение факторов.

3. Эффект мультиколлинеарности незначителен.

4. Коэффициенты модели значимы по  $t$ -критериям Стьюдента.

5. Модель адекватна по коэффициенту детерминации  $R^2$  и значима по  $F$ -критерию Фишера.

Целью данной работы является разработка общего алгоритма построения ХИК-моделей.

### Эффективный алгоритм "Selection B" организации "конкурса" моделей

Рассмотрим модель множественной линейной регрессии:

$$y_i = \alpha_0 + \alpha_1 x_{i1} + \alpha_2 x_{i2} + \dots + \alpha_m x_{im} + \varepsilon_i, \quad i = \overline{1, n}, \quad (1)$$

где  $y_i, i = \overline{1, n}$  – значения зависимой (объясняемой) переменной  $y$ ;

$x_{i1}, x_{i2}, \dots, x_{im}, i = \overline{1, n}$  – значения  $m$  независимых (объясняющих) переменных (регрессоров)  $x_1, x_2, \dots, x_m$ ;

$\varepsilon_i, i = \overline{1, n}$  – ошибки аппроксимации;

$\alpha_0, \alpha_1, \dots, \alpha_m$  – неизвестные параметры;

$n$  – объем выборки.

Рассмотрим изложенную, например, в работах [12, 13], технику оценивания неизвестных параметров модели (1) с помощью метода наименьших квадратов (МНК). Для этого проведем нормирование (стандартизацию) всех переменных по формулам:

$$v_i = \frac{y_i - \bar{y}}{\sigma_y}, \quad z_{i1} = \frac{x_{i1} - \bar{x}_1}{\sigma_{x_1}}, \quad \dots, \quad z_{im} = \frac{x_{im} - \bar{x}_m}{\sigma_{x_m}},$$

где  $\bar{y}, \bar{x}_1, \dots, \bar{x}_m$  – средние значения переменных;

$\sigma_y, \sigma_{x_1}, \dots, \sigma_{x_m}$  – среднеквадратические отклонения переменных;

$v, z_1, \dots, z_m$  – стандартизованные переменные, для которых среднее значение равно 0, а среднеквадратическое отклонение равно 1.

Тогда регрессии (1) ставится в соответствие её стандартизованная модель:

$$v_i = \beta_1 z_{i1} + \beta_2 z_{i2} + \dots + \beta_m z_{im} + u_i, \quad i = \overline{1, n}, \quad (2)$$

где  $\beta_1, \dots, \beta_m$  – неизвестные параметры (бэта-коэффициенты);

$u_i, i = \overline{1, n}$  – ошибки аппроксимации.

Уравнение стандартизованной регрессии в матричной форме имеет вид

$$v = Z\beta + u, \quad (3)$$

где  $v = \begin{bmatrix} v_1 \\ v_2 \\ \dots \\ v_n \end{bmatrix}$ ,  $Z = \begin{bmatrix} z_{11} & z_{12} & \dots & z_{1m} \\ z_{21} & z_{22} & \dots & z_{2m} \\ \dots & \dots & \dots & \dots \\ z_{n1} & z_{n2} & \dots & z_{nm} \end{bmatrix}$ ,  $\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \dots \\ \beta_m \end{bmatrix}$ ,  $u = \begin{bmatrix} u_1 \\ u_2 \\ \dots \\ u_n \end{bmatrix}$ .

Для оценки бета-коэффициентов с помощью МНК воспользуемся формулой:

$$\tilde{\beta}_{\text{МНК}} = (Z^T Z)^{-1} Z^T v. \quad (4)$$

При этом не трудно показать, что  $Z^T Z = r_{xx}$ , где  $r_{xx}$  – матрица коэффициентов парной корреляции между объясняющими переменными, а  $Z^T v = r_{yx}$ , где  $r_{yx}$  – вектор коэффициентов парной корреляции между объясняемой переменной  $y$  и объясняющими переменными  $x_1, x_2, \dots, x_m$ :

$$r_{yx} = \begin{bmatrix} r_{yx_1} \\ r_{yx_2} \\ \dots \\ r_{yx_m} \end{bmatrix}. \quad (5)$$

Тогда формула (4) для расчета бета-коэффициентов примет вид:

$$\tilde{\beta}_{\text{МНК}} = r_{xx}^{-1} r_{yx}. \quad (6)$$

Для перехода от бета-коэффициентов  $\tilde{\beta}_{\text{МНК}}$  к традиционным МНК-оценкам  $\tilde{\alpha}_{\text{МНК}}$  линейной регрессии (1) необходимо воспользоваться формулами [12]:

$$\tilde{\alpha}_i = \tilde{\beta}_i \frac{\sigma_y}{\sigma_{x_i}}, \quad i = \overline{1, m}, \quad (7)$$

$$\tilde{\alpha}_0 = \bar{y} - \tilde{\alpha}_1 \bar{x}_1 - \tilde{\alpha}_2 \bar{x}_2 - \dots - \tilde{\alpha}_m \bar{x}_m. \quad (8)$$

Коэффициент детерминации стандартизованной регрессии (2) в матричном виде находится по формуле [12]:

$$R^2 = r_{yx}^T \cdot \tilde{\beta}_{\text{МНК}}. \quad (9)$$

В работе [14] с использованием формул (6) и (9) автором был разработан новый алгоритм отбора информативных регрессоров (ОИР) в линейной регрессии по коэффициенту детерминации, который будем называть "Selection B". Его блок-схема представлена на рис. 1.

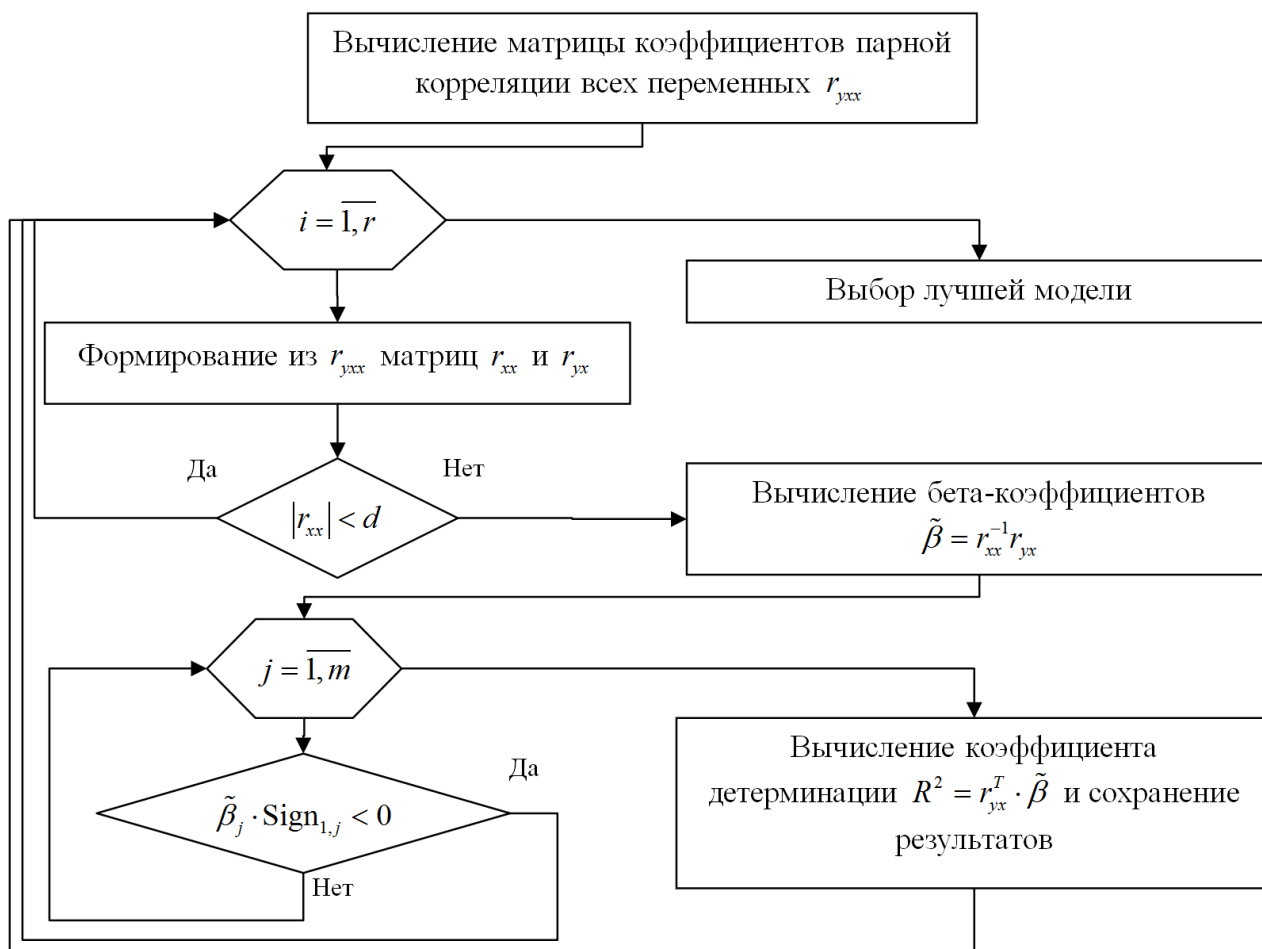


Рис. 1. Блок-схема алгоритма "Selection B"

На рис. 1  $\text{Sign}_{1 \times l}$  – вектор знаков влияния объясняющих переменных на  $y$ , который задается экспертами в исследуемой области по следующему правилу:

$$\text{Sign}_{1,j} = \begin{cases} -1, & \text{если } j\text{-я переменная оказывает отрицательное влияние на } y, \\ 1, & \text{если } j\text{-я переменная оказывает положительное влияние на } y, \\ 0, & \text{если затруднительно оценить влияние } j\text{-й переменной на } y. \end{cases}$$

Если хотя бы один знак бета-коэффициента  $\tilde{\beta}_j$  стандартизированной регрессии не согласуется с соответствующим знаком вектора  $\text{Sign}$ , т.е. выполняется условие  $\tilde{\beta}_j \cdot \text{Sign}_{1,j} < 0$ ,  $j = \overline{1, m}$ , то такая модель исключается из дальнейшего "конкурса".

Определитель матрицы парных коэффициентов корреляции  $|r_{xx}|$  представляет собой критерий обнаружения эффекта мультиколлинеарности. Если  $|r_{xx}| = 0$ , то имеет место совершенная мультиколлинеарность, а если  $|r_{xx}| = 1$ , то мультиколлинеарности нет. Если выполняется условие  $|r_{xx}| < d$ , где  $d$  – заданное пользователем пороговое значение, то такая регрессия сразу исключается из рассмотрения.

Достоинства алгоритма "Selection B":

- 1) время работы алгоритма не зависит от числа наблюдений  $n$ ;
- 2) возможность регулировать эффект мультиколлинеарности;

3) возможность контролировать физический смысл входящих в уравнение факторов.

Проведенный в работе [14] эксперимент показал, что проведение "конкурса" моделей с использованием пакета Gretl по алгоритму "Selection B" осуществляется более чем в 6 раз быстрее, чем в ПК АППРМ.

### **Фундаментальный блок алгоритмов построения ХИК-моделей**

Как отмечено выше, ХИК-регрессия должна удовлетворять пяти условиям: каждый коэффициент должен быть интерпретируем, знаки коэффициентов должны удовлетворять физическому смыслу факторов, эффект мультиколлинеарности должен быть незначителен, коэффициенты должны быть значимы по t-критерию Стьюдента и модель должна быть адекватна по коэффициенту детерминации и значима по F-критерию Фишера.

При использовании линейных регрессий вида (1) первое условие ХИК-моделей выполняется всегда. А при использовании, например, нелинейных аддитивных и линейно-мультипликативных регрессий оно справедливо не всегда. Таким образом, алгоритмы построения ХИК-моделей на основе различных форм связи между переменными естественным образом должны иметь свои индивидуальные особенности, поэтому в некоторой степени будут отличаться друг от друга, но при этом их основа, их ядро, "скелет", будет неизменным. Такой фундаментальный блок, образующий все алгоритмы построения ХИК-моделей, представлен на рис. 2. По нему видно, что для того чтобы очередная альтернативная регрессионная модель была допущена к участию в "конкурсе", она должна пройти жесткую процедуру отсева, состоящую из 5 стадий:

- 1) проверка "информативности" переменных [1];
- 2) проверка присутствия эффекта мультиколлинеарности;
- 3) проверка соответствия знаков коэффициентов физическому смыслу факторов;
- 4) проверка адекватности по коэффициенту детерминации и значимости в целом по F-критерию Фишера;
- 5) проверка значимости коэффициентов по t-критерию Стьюдента.

Если альтернатива не пройдет проверку хотя бы на одной стадии, то она выбывает из "конкурса". На практике может оказаться так, что ни одна альтернатива не пройдет проверку. В этом случае исследователю необходимо ослабить требования к модели.

Кратко охарактеризуем каждый из этапов блок-схемы, изображенной на рис. 2.

*Этап 1.* Вычисление матрицы парных коэффициентов корреляции. Этот этап является стандартным для всех алгоритмов и не представляет никаких проблем.

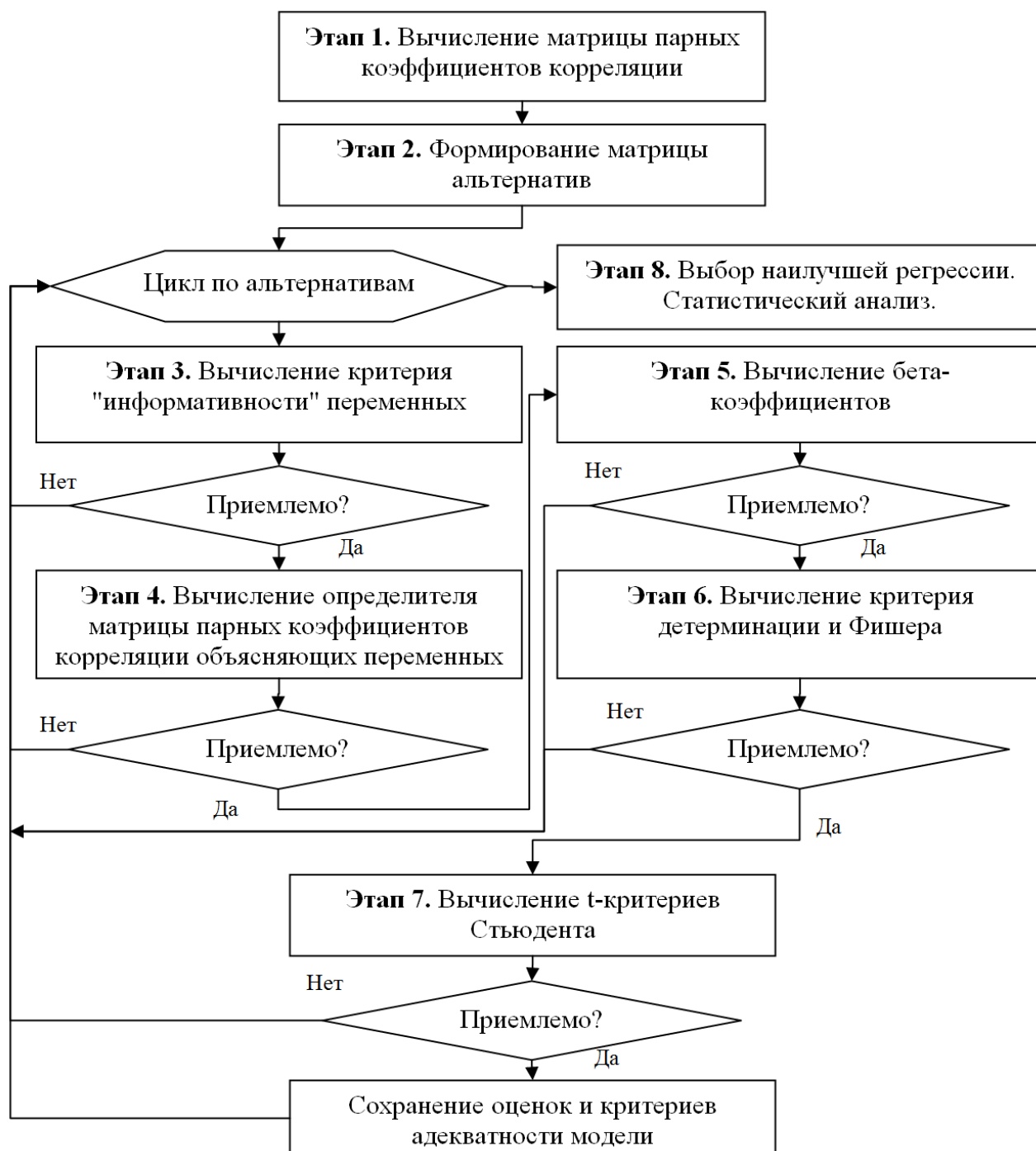
*Этап 2.* Формирование матрицы альтернатив. На этом этапе необходимо для заданной формы связи между переменными и стратегии её построения сформировать матрицу, содержащую построчно в некотором зашифрованном виде все возможные комбинации моделей.

*Этап 3.* Вычисление критерия "информативности" переменных. Подробное описание данного критерия приведено в монографии [1]. Для реализации этого этапа, во-первых, каждому регрессору предварительно должен быть поставлен в соответствие некоторый балл, например, по пятибалльной шкале. Во-вторых, должна быть назначена минимально возможная "информативность" переменных, которая позволит исключать из "конкурса" регрессии с низким значением критерия. Использование данного критерия для линейных регрессий не вызывает вопросов, а, например, для ЛМР необходимо решить вопрос с дублированием баллов в таких ситуациях.

*Этап 4.* Вычисление определителя матрицы парных коэффициентов корреляции объясняющих переменных. Напомним, что данный критерий обозначается  $|r_{xx}|$  и является показателем степени мультиколлинеарности. Данный этап является универсальным для линейных и квазилинейных регрессий.

*Этап 5.* Вычисление бета-коэффициентов. Бета-коэффициенты являются коэффициентами стандартизованной регрессии (2) и находятся по формуле (6). Вычисление этих коэффициентов необходимо для проверки соответствия их знаков содержательному

смыслу факторов. Поэтому предварительно, например, с привлечением экспертов в данной предметной области, необходимо сформировать вектор знаков влияния объясняющих переменных на  $y$  – вектор  $Sign_{1 \times k}$ . Условие проверки соответствия знаков для линейной регрессии приведено на блок-схеме алгоритма "Selection B" (рис. 1).



**Рис. 2.** Фундаментальный блок алгоритмов построения ХИК-моделей

*Этап 6.* Вычисление критерия детерминации и Фишера. Это стандартный для всех пяти форм этап. Для отсева незначимых моделей необходимо пользоваться таблицей критических точек распределения Фишера.

*Этап 7.* Вычисление t-критериев Стьюдента. Этот этап, как и предыдущий, является стандартным и для стандартизированной регрессии (2) реализуется по формулам:

$$t_i = \frac{\tilde{\beta}_i}{\sqrt{s^2 (r_{xx})_{ii}^{-1}}}, \quad i = \overline{1, m}, \quad (10)$$

где  $(r_{xx})_{ii}^{-1}$  –  $i$ -й диагональный элемент матрицы  $(r_{xx})^{-1}$ ;  $s^2 = \frac{1-R^2}{n-m-1}$  – величина остаточной дисперсии.

Для отсева незначимых моделей необходимо пользоваться таблицей критических точек распределения Фишера-Снедекора.

*Этап 8.* Выбор наилучшей регрессии. Статистический анализ. Данный этап предполагает работу с критериальной матрицей, содержащей только прошедшие многоэтапную процедуру жесткого отбора модели. Процедура выбора наилучшей регрессии для всех форм стандартна. Причем, выбор можно осуществлять либо по какому-то одному критерию адекватности, либо сразу по нескольким. Можно выбирать только одну лучшую регрессию, а можно, например, выбирать заданное число наилучших регрессий, т.е. формировать рейтинги моделей.

Для реализации алгоритма построения ХИК-моделей достаточно модернизировать эффективный алгоритм "Selection B", изображенный на рис. 1.

**Заключение.** В данной работе впервые введено понятие хорошо интерпретируемая качественная регрессионная модель (ХИК-модель). Сформулированы требования к таким регрессиям. Разработан фундаментальный блок алгоритмов построения ХИК-моделей. Основанные на этом блоке конкретные алгоритмы построения регрессионных моделей различных форм будут рассмотрены в будущих работах автора.

## БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Носков С.И. Технология моделирования объектов с нестабильным функционированием и неопределенностью в данных. – Иркутск: Облформпечать, 1996. – 321 с.
2. Носков С.И., Базилевский М.П. Программный комплекс автоматизации процесса построения регрессионных моделей // Международный журнал прикладных и фундаментальных исследований. – 2010. – № 1. – С. 93-94.
3. Базилевский М.П., Носков С.И. Методические и инструментальные средства построения некоторых типов регрессионных моделей // Системы. Методы. Технологии. – 2012. – № 1 (13). – С. 80-87.
4. Баенхаева А.В., Базилевский М.П., Носков С.И. Моделирование валового регионального продукта Иркутской области на основе применения методики множественного оценивания регрессионных параметров // Фундаментальные исследования. – 2016. – № 10-1. – С. 9–14.
5. Баенхаева А.В., Базилевский М.П., Носков С.И. Выбор структурной спецификации регрессионной модели валового регионального продукта Иркутской области // Информационные технологии и проблемы математического моделирования сложных систем. – 2016. – № 16. – С. 31-38.
6. Базилевский М.П., Врублевский И.П., Носков С.И., Яковчук И.С. Среднесрочное прогнозирование эксплуатационных показателей функционирования Красноярской железной дороги // Фундаментальные исследования. – 2016. – № 10-3. – С. 471–476.
7. Носков С.И., Врублевский И.П. Анализ регрессионной модели грузооборота железнодорожного транспорта // Вестник транспорта Поволжья. – 2020. – № 1 (79). – С. 86-90.

8. Носков С.И., Врублевский И.П. Регрессионная модель динамики эксплуатационных показателей функционирования железнодорожного транспорта // Современные технологии. Системный анализ. Моделирование. – 2016. – № 2 (50). – С. 192-197.

9. Базилевский М.П., Носков С.И. Моделирование обстановки с пожарами в сельских населенных пунктах в условиях их газификации // Информационные технологии и проблемы математического моделирования сложных систем. – 2012. – № 10. – С. 65-71.

10. Базилевский М.П., Носков С.И. Технология организации конкурса регрессионных моделей // Информационные технологии и проблемы математического моделирования сложных систем. – 2009. – № 7. – С. 77-84.

11. Базилевский М.П., Носков С.И. Статистический анализ критериальных матриц при организации "конкурса" регрессионных моделей // Информационные технологии и математическое моделирование в управлении сложными системами. – 2019. – № 1 (2). – С. 13-26.

12. Елисеева И.И. Эконометрика / И.И. Елисеева, С.В. Курышева, Т.В. Костеева и др. – М.: Финансы и статистика, 2007. – 576 с.

13. Фёрстер Э., Рёнц Б. Методы корреляционного и регрессионного анализа. – М.: Финансы и статистика, 1983. – 303 с.

14. Базилевский М.П. Повышение эффективности алгоритма отбора по критерию детерминации информативных регрессоров в регрессионных моделях // Прикладная математика и информатика: современные исследования в области естественных и технических наук : сборник научных статей IV научно-практической международной конференции (школы-семинара) молодых ученых: в двух частях. – 2018. – С. 196–202.

## REFERENCES

1. Noskov S.I. *Tehnologija modelirovanija ob'ektov s nestabil'nym funkcionirovanijem i neopredelennost'ju v dannyh* [Modeling technology for objects with unstable operation and data uncertainty]. Irkutsk, RIC GP «Oblinformpechat» Publ., 1996. 321 p.

2. Noskov S.I., Bazilevskiy M.P. *Programmnyj kompleks avtomatizacii processa postroenija regressionnyh modelej* [Program complex for automating the process of constructing regression models]. *Mezhdunarodnyj zhurnal prikladnyh i fundamental'nyh issledovanij* [International Journal of Applied and Basic Research]. 2010, no. 1, pp. 93–94.

3. Bazilevskiy M.P., Noskov S.I. *Metodicheskie i instrumental'nye sredstva postroenija nekotoryh tipov regressionnyh modelej* [Methodological and instrumental means of constructing some types of regression models]. *Sistemy. Metody. Tehnologii* [Systems. Methods. Technology]. 2012, no. 1, vol. 13, pp. 80–87.

4. Baenhaeva A.V., Bazilevskiy M.P., Noskov S.I. *Modelirovanie valovogo regional'nogo produkta Irkutskoj oblasti na osnove primeneniya metodiki mnozhestvennogo otsenivaniya regressionnykh parametrov* [Modeling of gross regional product Irkutsk region of the basis of methods of multiple estimation of regression parameters]. *Fundamental'nye issledovaniya* [Fundamental research]. 2016, no. 10-1, pp. 9–14.

5. Baenhaeva A.V., Bazilevskiy M.P., Noskov S.I. *Vybor strukturnoj specifikacii regressionnoj modeli valovogo regional'nogo produkta Irkutskoj oblasti* [The choice of the structural specification of the regression model of the gross regional product of the Irkutsk region]. *Informacionnye tehnologii i problemy matematicheskogo modelirovanija slozhnyh sistem* [Information technologies and problems of mathematical modeling of complex systems]. 2016, no. 16, pp. 31–38.

6. Bazilevskiy M.P., Vrublevskiy I.P., Noskov S.I., Yakovchuk I.S. *Srednesrochnoe prognozirovanie ekspluatatsionnykh pokazateley funkcionirovaniya Krasnoyarskoj zheleznoj dorogi* [Medium-term forecasting of performance indicators of functioning of Krasnoyarsk railway]. *Fundamental'nye issledovaniya* [Fundamental research]. 2016, no. 10-3. – pp. 471-476.



7. Noskov S.I., Vrublevskiy I.P. *Analiz regressionnoj modeli gruzooborota zheleznodorozhnogo transporta* [Analysis of the regression model of rail freight turnover]. *Vestnik transporta Povolzh'ja* [Volga Transport Bulletin]. 2020, no. 1, vol. 79, pp. 86–90.

8. Noskov S.I., Vrublevskiy I.P. *Regressionnaja model' dinamiki jekspluacionnyh pokazatelej funkcionirovanija zheleznodorozhnogo transporta* [Regression model of the dynamics of operational indicators of the functioning of railway transport]. *Sovremennye tehnologii. Sistemyj analiz. Modelirovanie* [Modern technologies. System analysis. Modeling]. 2016, no. 2, vol. 50, pp. 192–197.

9. Bazilevskiy M.P., Noskov S.I. *Modelirovanie obstanovki s pozharemi v sel'skih naseleennyh punktah v uslovijah ih gazifikacii* [Modeling the situation with fires in rural settlements in terms of their gasification]. *Informacionnye tehnologii i problemy matematicheskogo modelirovanija slozhnyh sistem* [Information technologies and problems of mathematical modeling of complex systems]. 2012, no. 10, pp. 65–71.

10. Bazilevskiy M.P., Noskov S.I. *Tehnologija organizacii konkursa regressionnyh modelej* [Technology for organizing a competition of regression models]. *Informacionnye tehnologii i problemy matematicheskogo modelirovanija slozhnyh sistem* [Information technologies and problems of mathematical modeling of complex systems]. 2009, no. 7, pp. 77–84.

11. Bazilevskiy M.P., Noskov S.I. *Statisticheskij analiz kriterial'nyh matric pri organizacii "konkursa" regressionnyh modelej* [Statistical analysis of criterion matrices when organizing a "competition" of regression models]. *Informacionnye tehnologii i matematicheskoe modelirovanie v upravlenii slozhnymi sistemami* [Information technology and mathematical modeling in the management of complex systems]. 2019, no. 1, vol. 2, pp. 13–26.

12. Eliseeva I.I., Kurysheva S.V., Kosteeva T.V., etc. *Jekonometrika* [Econometrics]. Moscow, Finance and Statistics, 2007. 576 p.

13. Ferster A., Renz B. *Metody korreljacionnogo i regressionnogo analiza* [Correlation and regression analysis methods]. Moscow, Finance and Statistics, 1983. 303 p.

14. Bazilevskiy M.P. *Povyshenie jeffektivnosti algoritma otbora po kriteriju determinacii informativnyh regressorov v regressionnyh modeljah* [Improving the efficiency of the selection algorithm by the criterion of determination of informative regressors in regression models]. *Prikladnaja matematika i informatika: sovremennye issledovanija v oblasti estestvennyh i tehniceskikh nauk : sbornik nauchnyh statej IV nauchno-prakticheskoy mezhdunarodnoj konferencii (shkoly-seminara) molodyh uchenyh: v dvuh chastjah* [Applied mathematics and computer science: modern research in the field of natural and technical sciences: collection of scientific articles of the IV scientific and practical international conference (school-seminar) of young scientists: in two parts]. 2018, pp. 196–202.

### Информация об авторах

*Михаил Павлович Базилевский* – к. т. н., доцент, доцент кафедры «Математика», Иркутский государственный университет путей сообщения, г. Иркутск, e-mail: mik2178@yandex.ru

### Authors

*Mikhail Pavlovich Bazilevskiy* – Ph. D. in Engineering Science, Associate Professor, the Subdepartment of Mathematics, Irkutsk State Transport University, Irkutsk, e-mail: mik2178@yandex.ru

### Для цитирования

Базилевский М.П. Фундаментальный блок алгоритмов построения хорошо интерпретируемых качественных регрессионных моделей // «Информационные технологии и математическое моделирование в управлении сложными системами»: электрон. науч. журн. – 2020. – №3(8). – С. 1-10 – DOI: 10.26731/2658-3704.2020.3(8).1-10 – Режим доступа:

**For citations**

Bazilevskiy M.P. The fundamental block of algorithms for constructing well-interpreted qualitative regression models // *Informacionnye tehnologii i matematicheskoe modelirovanie v upravlenii slozhnymi sistemami: ehlektronnyj nauchnyj zhurnal* [Information technology and mathematical modeling in the management of complex systems: electronic scientific journal], 2020. No. 3(8). P. 1-10. DOI: 10.26731/2658-3704.2020.3(8).1-10 [Accessed 01/11/20]