

УДК 519.6

Г.Р. Кадырова

ОЦЕНКА И ПРОГНОЗИРОВАНИЕ СОСТОЯНИЯ ТЕХНИЧЕСКОГО ОБЪЕКТА ПО РЕГРЕССИОННЫМ МОДЕЛЯМ

Кадырова Гульнара Ривальевна, кандидат технических наук, окончила радиотехнический факультет Ульяновского политехнического института. Доцент кафедры «Прикладная математика и информатика» УлГТУ. Имеет статьи, монографии, учебные пособия в области статистического моделирования, программных информационных систем. [e-mail: gulya@ulstu.ru].

Аннотация

В статье представлен статистический пакет «Система поиска оптимальных регрессий» (СПОР), реализующий подход адаптивного регрессионного моделирования и обеспечивающий оценку адекватности модели наблюдениям и поиск ее оптимальной структуры. Эффективность данной методологии, под которой понимается сокращение размерности модели и повышение точности определения ее параметров и прогноза, прямо пропорциональна размерности, степени зашумленности и мультиколлинеарности исходных данных, что позволяет считать ее применение для описания состояния технических объектов перспективным математическим подходом.

Одной из задач при анализе данных является задача выбора меры сравнения конкурирующих моделей. Для моделей небольшой размерности, предназначенных для прогноза, использование внутренних мер качества не всегда может дать точное представление о предпочтительности той или иной конкурирующей структуры.

Исследованы свойства меры «скользящего экзамена», основанной на всей выборке данных и использующей ее в качестве контрольной, по отношению к известным внутренним и внешним мерам. Показана перспектива ее применения для идентификации оптимальной модели прогноза в рамках СПОР.

Ключевые слова: регрессионное моделирование, прогнозирование, методы структурной идентификации, меры качества, система поиска оптимальных регрессий, статистический пакет.

EVALUATION AND PREDICTION OF TECHNICAL OBJECT CONDITION WITH THE USE OF REGRESSION MODELS

Gulnara Rivalevna Kadyrova, Candidate of Engineering; graduated from the Faculty of Radioengineering of Ulyanovsk Polytechnic Institute; Associate Professor of the Department of Applied Mathematics and Informatics of Ulyanovsk State Technical University; an author of monographs, textbooks, and articles in the field of statistical modeling, software information systems. e-mail: gulya@ulstu.ru.

Abstract

The article presents information about the System of Optimum Regression Search (SORS) statistical package realizing the approach of adaptive regression modelling and providing an evaluation of the observation model adequacy and search of its optimal structure. The methodology means reduction of the model dimension and increase of determination accuracy of its parameters and the forecast. Efficiency of the methodology is directly proportional to dimension, the degree of noisiness, and multicollinearity of the initial data. Consequently, it allows considering its application for the description of technical objects conditions as perspective mathematical approach.

One of the tasks in data analysis is the problem of choice of a comparison measure for the competitive models. The use of internal quality measures for smaller dimension models intended for forecasting cannot always give as authentic view of a particular competitive structure that is preferable as possible.

The article investigates properties of the cross validation measure that is based on the complete sample of data and uses it as control data sample for familiar internal and external measures. The prospect of its application for identification of the forecasting optimal model within SORS is demonstrated in the article.

Key words: regression modelling, forecasting, methods of structural identification, quality measure, SORS, statistical package.

ВВЕДЕНИЕ

Получение моделей, адекватно описывающих состояние технических объектов с последующим их использованием для прогноза выходных характеристик (откликов), является весьма актуальной задачей. Поиск адекватной модели в рамках подхода регрессионного моделирования (РМ) [1] является чрезвычайно трудозатратной процедурой, требующей использования в процессе адаптации набора критериев, базисных функций, методов структурно-параметрической идентификации, а также возможности реализации различных логически обоснованных сценариев обработки.

Разработанный программный пакет СПОР (система поиска оптимальных регрессий) [2] является специализированной системой, реализующей стратегию адаптивного РМ. Пакет обеспечивает оценку адекватности модели наблюдениям и поиск ее оптимальной структуры, эффективную настройку на различные сценарии обработки и автоматизацию процесса адаптации вычислительных схем к нарушениям условий метода наименьших квадратов (МНК) [3].

Одной из важнейших задач при анализе данных является задача выбора меры сравнения конкурирующих описаний. Для моделей небольшой размерности, предназначенных для прогноза, использование внутренних мер качества, например средней квадратической ошибки, не всегда может дать точное представление о предпочтительности той или иной конкурирующей структуры. В ходе надежного вычислительного эксперимента удалось показать, что наилучшей может быть мера, построенная на принципе «скользящего экзамена» (СЭ). Ее применение гарантирует при заданной доверительной вероятности попадание ошибки прогноза в оптимальный интервал по внешним критериям точности.

1 АДАПТИВНОЕ РМ

Альтернативой традиционному подходу является стратегия адаптивного РМ, нацеленная на оптимизацию модели обработки данных и предусматривающая на этапе получения МНК-оценок параметров статистический анализ модели и ее членов (регрессионный анализ – РА), а также использующая дополнительные этапы обработки: проверку соблюдения предположений РА-МНК, адаптацию в случае их нарушений. Регрессионное моделирование – это системный подход, при котором корректность применения любого элемента системы (выборки, модели, методов оценивания параметров и структурной идентификации, критерия качества, набора предположений) может быть подвергнута сомнению и проверке с соответствующей адаптацией при нарушении заданных условий.

Запишем исходную модель РА в матричном виде:

$$Y = X\beta + \varepsilon, \quad (1)$$

где Y – вектор наблюдений размера $(n \times 1)$, содержащий n наблюдаемых значений y_1, y_2, \dots, y_n зависимой переменной или отклика Y (Y – случайная величина);

X – регрессионная матрица размера $(n \times p)$, содер-

жащая $x_{10} = x_{20} = \dots = x_{n0} = 1$ и элементы x_{ij} как i -е наблюдения ($i = \overline{1, n}$) над регрессорами x_{ij} ($j = \overline{1, p-1}$) (x_{ij} – неслучайная величина), p – количество регрессоров;

β – вектор размера $(p \times 1)$ истинных неизвестных параметров β_j ($j = \overline{0, p-1}$), подлежащих оцениванию (β_j – неслучайная величина);

ε – вектор ошибок размера $(n \times 1)$, содержащий неизвестные погрешности наблюдений ε_i ($i = \overline{1, n}$) (ε_i – случайная величина).

В линейном РА модель (1) считается неизвестной как по структуре, под которой понимается состав и размерность модели, так и по параметрам – элементам β_j – вектора β . Назовем задачей параметрического оценивания задачу нахождения по результатам наблюдений оптимальных оценок β_j . Задачу нахождения оптимальной структуры модели можно назвать тогда задачей структурной идентификации.

Подход адаптивного РМ предлагает проверку соблюдения условий применения МНК: постоянство дисперсии, независимость регрессоров, нормальность распределения ошибок, нулевое значение математического ожидания ошибок, независимость ошибок. При этом избыточные члены модели находят с помощью t -критерия; мультиколлинеарность может быть обнаружена по коэффициентам парной корреляции. С целью максимального ослабления влияния взаимозависимости регрессоров используются центрирование данных и далее устойчивые к мультиколлинеарности вычислительные схемы МНК (метод Хаусхолдера). Адаптация к нарушению предположения о нормальности выполняется применением робастных методов. Нарушение условия однородности наблюдений или постоянства дисперсии проверяется по графикам остатков. Для проверки нарушения условия независимости ошибок используется критерий Дарбина-Уотсона. Если основные условия применения МНК соблюдаются, построенная модель может использоваться для прогнозирования.

Эффективность данной методологии, под которой понимается сокращение размерности модели и повышение точности определения ее параметров и прогноза, прямо пропорциональна размерности, степени зашумленности и мультиколлинеарности исходных данных, что позволяет считать ее применение для описания состояния технических объектов наиболее перспективным математическим подходом.

2 ПРОГРАММНЫЙ ПАКЕТ СПОР

При организации оптимальной предметной стратегии РМ необходимо учесть наличие различных выборок, предположений, классов функций, методов оценивания, мер качества и их наборов для принципа многокритериальности, методов структурной идентификации, конкурирующих в соответствии с принципом неокончателных решений стратегий адаптации к нарушению предположений.

Для практического применения РМ прежде всего требуется полная автоматизация всех заявленных процедур,

для чего и было разработано соответствующее программное обеспечение [4].

При разработке программного пакета СПОР ставились и решались проблемы:

- автоматизации обработки информационных потоков и ведения интерактивного взаимодействия при реализации сценария поиска оптимальной модели;
- наполнения статистической подсистемы;
- обработки малых выборок (проблемы внешних мер);
- обработки больших выборок (проблемы выбора метода поиска);
- реализации сценариев для различных классов моделей и др.

Пакет СПОР включает в себя следующие модули [5]:

- 1) управляющий модуль;
- 2) модуль формирования запроса;
- 3) библиотеку функциональных процедур;
- 4) блок сценария;
- 5) блок настройки системы;
- 6) блок редактора данных;
- 7) блок формирования таблиц;
- 8) справочник.

Сочетание такой структуры с системой меню и диалоговым режимом работы делает работу в СПОР очень удобным.

Основным инструментом положительного воздействия на прогностические свойства модели является алгоритм поиска ее оптимальной структуры.

В пакете реализованы следующие процедуры структурно-параметрической идентификации:

- множественная линейная регрессия,
- гребневая регрессия,
- робастное оценивание,
- полный перебор структур,
- неполный перебор структур (перебор с ограничением на количество включаемых регрессоров в модель),
- перебор нормальных систем,
- пошаговая регрессия с включением-исключением,
- случайный поиск с адаптацией,
- случайный поиск с возвратом.

Следует отметить, что данные процедуры могут быть вызваны как в автоматическом режиме для обработки целого ряда выборок данных, так и при обработке отдельной выборки данных по реализованному оптимальному сценарию.

В пакете реализована процедура построения и анализа графика остатков, что является полезным статистическим инструментом для проверки адекватности оцененной модели регрессии имеющимся данным.

Конкурентоспособность СПОР с другими статистическими пакетами:

- использование новых методов структурной идентификации: полный перебор, неполный перебор переопределенных и нормальных систем, многокритериальный метод пошаговой регрессии с включением-исключением;
- использование гибкого инструмента построения сравнительных таблиц;

- использование, помимо классических (внутренних) мер качества, внешних мер и мер СЭ, что позволяет осуществлять оценку внешней адекватности модели (точности прогноза).

Первая версия программного пакета была реализована на языке Object Pascal с использованием библиотек Turbo Vision и функционировала под управлением ОС MS DOS. Версия СПОР 2.0. [2] разработана на платформе ОС Windows.

В настоящее время ведутся работы по наращиванию функциональных возможностей СПОР и ее интеллектуализации [6, 7].

3 МЕРЫ КАЧЕСТВА МОДЕЛИ

Одной из важнейших задач при анализе данных является задача выбора меры сравнения конкурирующих описаний.

В рамках СПОР предоставляется возможность, помимо классических внутренних мер, вычисления внешних мер и мер СЭ.

3.1 Внутренние меры качества

Качество модели РА обычно определяют по внутренним мерам – средней квадратической ошибке σ , которая применяется как для оценки адекватности модели, так и для сравнения различных моделей друг с другом; выборочному коэффициенту множественной корреляции R , который используют как меру линейной связи (1): чем больше значение R ($0 \leq R \leq 1$), тем сильнее связь, то есть тем лучше аппроксимирующая функция соответствует наблюдениям, также высокое значение R гарантирует пригодность модели для прогноза; F – критерию, при $F > 4F_T$ ($\alpha; p - 1, n - 1$) (F_T – критическое значение, взятое из таблицы для F -критерия) модель признается заслуживающей внимания на предмет ее использования для прогноза.

Оценив эффективность модели по одной или нескольким мерам и признав ее пригодной к использованию, затруднительно полагать ее полностью адекватной – такой, что никакой другой набор регрессоров из использованных не даст статистически значимого улучшения в прогнозе \hat{Y} . Задача формирования полностью адекватной модели требует анализа каждого регрессора, включенного в модель.

Проверка гипотезы $H_0: \beta_j = 0$ для $j = \overline{1, p - 1}$ означает проверку предположения о том, что регрессор x_j не улучшает прогноз \hat{Y} по сравнению с прогнозом, получаемым с помощью модели по $(p - 2)$ другим регрессорам, то есть проверку предположения о том, что x_j – статистически незначимая переменная. Наиболее просто гипотезу можно проверить, применяя статистику

$$F_j = \frac{\hat{\beta}_j^2}{\hat{\sigma}_j^2}, \quad (2)$$

где $\hat{\sigma}_j^2$ – диагональные элементы дисперсионно-ковариационной матрицы $D\hat{\beta}$. В отличие от общего

F -критерия, статистика F_j называется частным F -критерием. Для H_0 эта статистика имеет F -распределение с 1 и $(n-p)$ степенями свободы. Кроме (2), часто используется статистика

$$t_j = \frac{\hat{\beta}_j}{\hat{\sigma}_j}, \quad (3)$$

имеющая для H_0 t -распределение с $(n-p)$ степенями свободы. Значению t_j удовлетворяющему неравенству $t_j \leq t_T$, при котором гипотеза H_0 отвергается на уровне значимости α , соответствует согласно формуле (3) значение $\hat{\sigma}_j$, равное определяемой оценке $\hat{\beta}_j$ или меньше ее.

Мера t является полностью структурной мерой. С ее помощью решается задача выбора информативного по t -критерию набора регрессоров. Мера необязательно приводит к набору с минимальной ошибкой, к тому же эффективность ее зависит от степени соблюдения предположения <3.1> [1] о независимости регрессоров.

Итак, при статистическом анализе регрессоров в первую очередь оценивается значимость каждого члена модели. Кроме этого, возможно еще оценить и степень важности каждого регрессора, анализируя так называемые частные коэффициенты корреляции.

В результате статистического анализа регрессоров можно установить их значимость, а также проранжировать по степени важности для регрессионной модели.

Рассмотренные внутренние меры качества характеризуют адекватность модели только по отношению к использованной для ее построения выборке точек. Это первый этап исследования модели, на котором экспериментатор должен быть убежден, что модель соответствует наблюдениям.

3.2 Внешние меры качества

Если модель предназначена для прогноза, то надо быть уверенным в ее пригодности для определения области, не совпадающей с выборочными точками y_i .

Для оценки внешней адекватности (точности прогноза) используются контрольные точки. Общим для всех подходов к построению внешних мер является деление исходной выборки на обучающую (ОВ) и контрольную (КВ). На первой выборке строится модель или множество моделей; на второй – выполняется оценка ее адекватности или дискриминация по статистикам – внешним мерам качества.

Внешние меры основаны на анализе расхождений между прогнозом \hat{Y} и известным наблюдаемым значением Y для объектов, не участвовавших в получении модели.

Назовем разность между наблюдаемым значением отклика y_i для объекта, не использованного в модели для ОВ, и его прогнозом \hat{y}_i расхождением Δ_i в отличие от понятия остатка e_i для точек ОВ. Расхождение Δ_i можно записать в виде:

$$\Delta_i = y_i - \hat{y}_i, \quad (4)$$

где $i = \overline{1, n_{KB}}$ и n_{KB} – количество точек в КВ.

Мера

$$\bar{\Delta} = \sum_{i=1}^{n_{KB}} \frac{\Delta_i}{n_{KB}} \quad (5)$$

или ее абсолютная величина

$$|\bar{\Delta}| = \sum_{i=1}^{n_{KB}} \frac{|\Delta_i|}{n_{KB}} \quad (6)$$

могут служить характеристикой систематической ошибки, обусловленной в общем смысле методом обработки, а статистики

$$\sigma_{\Delta_1} = \sqrt{\frac{\sum_{i=1}^{n_{KB}} \Delta_i^2}{n_{KB} - p}}, \quad (7)$$

$$\sigma_{\Delta_2} = \sqrt{\frac{\sum_{i=1}^{n_{KB}} (\Delta_i - \bar{\Delta})^2}{n_{KB} - p}}, \quad (8)$$

$$\sigma_{\Delta_{n-1}} = \sqrt{\frac{\sum_{i=1}^{n_{KB}} \Delta_i^2}{n_{KB} - 1}}, \quad (9)$$

$$\sigma_{|\Delta|} = \sqrt{\sum_{i=1}^n \frac{(|\Delta_i| - \bar{\Delta})^2}{n_{KB} - p}} \quad (10)$$

– характеристиками случайной выборки.

Мера R_{Δ} может рассматриваться как мера линейной связи между прогнозом и наблюдением: чем выше R_{Δ} , тем лучше модель соответствует наблюдениям.

Для оценки качества прогноза можно использовать так называемый средний квадрат ошибки предсказания (СКОП) [8] или безусловный средний квадрат ошибки UMSE [9]:

$$СКОП = \frac{(n_{KB}^2 - n_{KB} - 2)(n_{KB} - 1)\hat{\sigma}_{\Delta}^2(1 - \hat{R}_{\Delta}^2)}{n_{KB}(n_{KB} - p)(n_{KB} - p - 1)}. \quad (11)$$

3.3 Меры СЭ

Поскольку, работая с малыми выборками, нет возможности разделить ее на обучающую и контрольную с достаточно большим количеством точек, для оценки внешней адекватности мы предлагаем использовать внешние меры, основанные на «скользящей» контрольной выборке [10]. Если последовательно каждый из объектов выборки выводить из нее, полагая этот объект контрольным, и пересчитывать заново параметры модели, то разности между y_i и \hat{y}_i для скользящей контрольной точки $\Delta_i =$ «наблюдение минус прогноз» ($i = \overline{1, n}$; где n – общее количество объектов) могут быть использованы для построения ряда внешних мер, в дальнейшем для удобства называемых мерами СЭ.

Последовательное исключение объектов, соответствующее удалению определенных строк из матрицы данных X , дает возможность сформулировать искусственно новую выборку (проверочную или контрольную) того же объема, что и исходная.

Внешние меры, основанные на поединочном выделении всех объектов в группу контрольных, позволяют

исследовать модель на всем исходном выборочном пространстве, в отдельных зонах и точках.

Введем понятие скользящего расхождения (c -расхождения), имея в виду в (4) под $y_i (i = \overline{1, n})$ значения наблюдаемых откликов для объектов выборки, а под \hat{y}_i – прогнозы по моделям, полученным по выборкам объема $(n - 1)$, из которых каждый раз при образовании Δ_i удаляется i -й объект. Таким образом, имеющаяся в распоряжении исследователя выборка используется одновременно и как обучающая, и как контрольная, точнее, как скользящая контрольная выборка. Выражение (4) можно переписать в виде:

$$\Delta_i = y_i - x_i^T \hat{\beta}_c, \quad (12)$$

где x_i^T – i -я строка матрицы X ;

$\hat{\beta}_c$ – оценка вектора $\hat{\beta}$, полученная с матрицей X без этой i -й строки и вектором Y без i -го элемента в соответствии с формулой $\hat{\beta} = (X^T X)^{-1} X^T Y$.

Соблюдаемое обычно на практике неравенство $\Delta_i > e_i$ является следствием того, что $\hat{\beta}_c \neq \hat{\beta}$. Модель $\hat{Y} = X \hat{\beta}_c$ для объема выборки $(n - 1)$ «не схватывает» в отличие от $\hat{Y} = X \hat{\beta}$ для объема n индивидуальные особенности выделенной из выборки i -й точки, увеличивая тем самым различие между наблюдением и прогнозом. Следовательно, получаемые таким образом c -расхождения будут оценивать качество прогноза в точках зон, являющихся окрестностями точек выборки. Естественно, в последних качество прогноза определяется остатками e_i . Оценки c -расхождений будут несколько завышены по сравнению с реальными расхождениями, так как в действительности для прогноза используется модель по выборке объема n .

Таким образом, с помощью c -расхождений, построенных на точках выборки, мы характеризуем прогностические свойства соответствующей ей модели в окрестностях этих точек. При этом предполагается, что модель для полной выборки и n моделей скользящей процедуры существенно не отличаются друг от друга, т. е. аномальные выбросы должны быть предварительно выведены из выборки.

Если в РА везде остатки e_i заменить на c -расхождения Δ_i , то на всех стадиях статистического анализа можно получить более объективные суждения, чем ранее. При этом все внутренние критерии превратятся во внешние; по ним будет оцениваться адекватность модели на скользящей выборке (псевдовыборке), играющей роль контрольной. Проверка предпосылок <1>–<5> [1] по Δ_i и критериям, основанным на c -расхождениях, обеспечивает относительно точные результаты и быструю сходимость процесса последовательной адаптации в РМ.

3.4 Численный анализ мер

Исследовались свойства меры СЭ. Решалась задача численного исследования свойств меры СЭ по отношению к внутренним и внешним мерам, перспективы ее примене-

ния для идентификации оптимальных моделей.

Анализ мер выполнен пакетом СПОР по результатам обработки 45 выборок данных, предоставленных Казанским государственным университетом.

Статистический анализ моделей для 45 выборок позволил выяснить, какая мера: стандартная ошибка σ или мера СЭ $\sigma_{СЭ}$ – лучше приближается к внешней мере σ_{Δ} .

Сравнение по точности прогноза моделей, полученных по внутренней мере σ и мере СЭ $\sigma_{СЭ}$, проводилось по отношению дисперсий: $\frac{(\sigma_{\Delta}^{\sigma})^2}{(\sigma_{\Delta}^{\sigma})^2}$ и $\frac{(\sigma_{\Delta}^{СЭ})^2}{(\sigma_{\Delta}^{\sigma})^2}$ и разностей

случайных ошибок: $\sigma_{\Delta}^{\sigma} - \sigma_{\Delta}^{\Delta}$ и $\sigma_{\Delta}^{СЭ} - \sigma_{\Delta}^{\Delta}$, где σ_{Δ}^{σ} , $\sigma_{\Delta}^{СЭ}$ – внешняя мера σ_{Δ} для оптимальных моделей, полученных по мерам σ и $\sigma_{СЭ}$ соответственно; σ_{Δ}^{Δ} – внешняя мера σ_{Δ} для оптимальных моделей, полученных по мере σ_{Δ} .

Результаты сравнения отражают явное преимущество меры СЭ по сравнению с внутренней мерой: в 30 случаях зарегистрировано значимое (отношение дисперсий больше 2,18) ухудшение прогностических свойств моделей, полученных по мере σ по сравнению с 12 случаями при СЭ. Если весь диапазон полученных значений отношений дисперсий разделить на интервалы, получим, что при СЭ основная доля (11 случаев) попадает в интервал [2; 4] и лишь в одном случае зарегистрировано ухудшение прогностических свойств в 5,623 раза; в десяти случаях – совпадение по точности прогноза, в остальных случаях – незначимое ухудшение прогноза (диапазон отношений дисперсий: [1; 2]). Для моделей, полученных по внутренней мере, основная доля (22 случая) значимого ухудшения прогностических свойств попадает в интервал [3; 14]; максимальное ухудшение, равное 13,171, зарегистрировано для выборки № 339; полное совпадение по точности прогноза получено только в четырех случаях. Таким образом, модели, полученные по мере СЭ, по прогностическим свойствам примерно в 3 раза лучше моделей, полученных по внутренним мерам.

Кроме улучшения по точности прогноза, модели, полученные по мере $\sigma_{СЭ}$, почти не содержат незначимых по t -статистике регрессоров. В 70% всех случаев модели, оптимальные по внутренней мере, содержат малоинформативные слагаемые. Модели, оптимальные по мере СЭ, лишь в 17% всех случаев содержат незначимые слагаемые. Модели, содержащие малоинформативные слагаемые, полученные по мере СЭ, содержат по одному незначимому регрессору, в то время как модели, полученные по внутренней мере σ , – как правило, два и более малоинформативных слагаемых.

Было оценено, значимое ли улучшение по внешней точности дает использование в качестве критерия качества меры СЭ. Для этого вычислялось отношение средних

дисперсий $\frac{(\bar{\sigma}_{\Delta}^{\sigma})^2}{(\bar{\sigma}_{\Delta}^{СЭ})^2}$ и сравнивалось с табличным значени-

ем F_T . Анализ показал, что использование меры СЭ дает значимое улучшение прогностических свойств по сравнению с внутренними мерами.

ЗАКЛЮЧЕНИЕ

Использование статистического пакета СПОР для оценки и прогнозирования состояния технического объекта позволяет осуществлять эффективный поиск адекватной модели за счет автоматизации процесса адаптации вычислительных схем к нарушениям условий применения РА–МНК, значительно сокращая время по сравнению с применением автономно работающих процедур.

Путем вычислительного эксперимента показано, что модели, полученные по мерам СЭ, по прогностическим свойствам лучше моделей, полученных по внутренним мерам, то есть меры СЭ лучше приближаются к внешним мерам. Помимо этого меры СЭ обеспечивают устранение незначимых слагаемых (сокращение размерности модели).

Таким образом, если модель предназначена для прогноза, то для малых выборок в качестве меры поиска оптимальной модели лучше использовать меры СЭ.

СПИСОК ЛИТЕРАТУРЫ

1. Валеев С.Г. Регрессионное моделирование при обработке наблюдений. – М.: Наука, 1991. – 272 с.
2. Валеев Г.Р., Кадырова Г.Р. Система поиска оптимальных регрессий. – Казань: ФЭН, 2003. – 160 с.
3. Валеев Г.Р., Кадырова Г.Р. Автоматизированная система для решения задач метода наименьших квадратов // Известия Вузов. Сер. Геодезия и аэрофотосъемка. – 1999. – № 6. – С. 9–14.
4. Валеев С.Г., Кадырова Г.Р., Турченко А.А. Программная система поиска оптимальных регрессий // Вопросы современной науки и практики. Сер. Технические науки. – 2008. – № 4 (14), т. 2. – С. 97–101.
5. Кадырова Г.Р. Программная система поиска оптимальных регрессионных моделей прогноза // Путь науки. – 2014. – № 7 (7). – С. 10–11.
6. Кадырова Г.Р., Турченко А.А., Касимова Л.Л. Программная реализация переборных алгоритмов структурной идентификации на основе многопоточного программирования // Информационные и интеллектуальные технологии: тр. междунар. конф. по логике, информатике, науковедению. – Ульяновск, 2007. – С. 34–36.
7. Кадырова Г.Р. Система поиска оптимальной модели. Состояние дел и перспективы развития // Потенциал современной науки. – 2015. – № 4 (12). – С. 8–10.
8. Себер Д. Линейный регрессионный анализ. – М.: Мир, 1980. – 450 с.
9. Афифи А., Эйзен С. Статистический анализ. Подход с использованием ЭВМ. – М.: Мир, 1982. – 487 с.
10. Kadyrova G.R. About sliding exam kriteria in the regression tasks of small dimensionality // Abstracts of

papers submitted to the Russian-American Microsymposium on Planetology. Vernadsky Institute of geochemistry and analytical chemistry Russian Academy of Sciences. – 1997. – pp. 47–48.

REFERENCES

1. Valeev S.G. *Regressiionnoe modelirovanie pri obrabotke nabludeni* [Regression Modeling at Reducing of Observations]. Moscow, Nauka Publ., 1991. 272 p.
2. Valeev G.R., Kadyrova G.R. *Sistema poiska optimalnykh regressii* [The System for Optimal Regressions Search]. Kazan, FJEN Publ., 2003. 160 p.
3. Valeev G.R., Kadyrova G.R. *Avtomatizirovannaia sistema dlia resheniia zadach metoda naimenshih kvadratov* [Computer-Aided System for Solving the Problem of the Least Square Method]. *Izvestiia Vuzov. Ser.: Geodeziia i aerofotosemka* [Geodesy and Aerophotography], 1999, no. 6, pp. 9–14.
4. Valeev S.G., Kadyrova G.R., Turchenko A.A. *Programmnaia sistema poiska optimalnykh regressii* [Software for Optimal Regressions Search]. *Voprosy sovremennoi nauki i praktiki. Ser. Tekhnicheskie nauki* [Problems of Contemporary Science and Practice Journal. Technical Sciences Series], 2008, no. 4 (14), vol. 2, pp. 97–101.
5. Kadyrova G.R. *Programmnaia sistema poiska optimalnykh regressiionnykh modelei prognoza* [Program Search System of Optimal Regression Prediction Models]. *Put nauki* [The Way of Science], 2014, no. 7 (7), pp. 10–11.
6. Kadyrova G.R., Turchenko A.A., Kasimova L.L. *Programmnaia realizatsiia perebornykh algoritmov strukturnoi identifikatsii na osnove mnogopotchnogo programmirovaniia* [Software Implementation of Search Algorithms of the Identification by Diffraction Analysis on the Basis of the Multithreaded Programming]. *Informatsionnye i intellektualnye tekhnologii: tr. mezhdunar. konf. po logike, informatike, naukovedeniiu* [Information and Intelligent Technologies: Proc. of the Intern. Conf. on Logic, Informatics, and Science of Science]. Ulyanovsk, 2007, pp. 34–36.
7. Kadyrova G.R. *Sistema poiska optimalnoi modeli. Sostoianie del i perspektivy razvitiia* [The Search System of the Optimal Model. State and Prospects of Development]. *Potentsial sovremennoi nauki* [The Potential of Modern Science], 2015, no. 4 (12), pp. 8–10.
8. Seber D. *Lineinyi regressiionnyi analiz* [Linear Regression Analysis]. Moscow, Mir Publ., 1980. 450 p.
9. Afifi A., Azen S. *Statisticheskii analiz. Podhod s ispolzovaniem EVM* [Statistical Analysis. A Computer Oriented Approach]. Moscow, Mir Publ., 1982. 487 p.
10. Kadyrova G.R. About Sliding Exam Kriteria in the Regression Tasks of Small Dimensionality. *The Russian-American Micro Symposium on Planetology*. Vernadsky Institute of Geochemistry and Analytical Chemistry of RAS, 1997, pp. 47–48.