УДК 519.233.5 DOI 10.18522/2311-3103-2022-5-152-160

В.Н. Лутай, Н.Ш. Хусаинов

ПОВЫШЕНИЕ УСТОЙЧИВОСТИ ЛИНЕЙНОЙ РЕГРЕССИИ

Рассматривается метод повышения устойчивости линейного регрессионного уравнения, коэффициенты которого являются решением системы нормальных уравнений. Неустойчивость уравнения определяется наличием линейной зависимости между экспериментальными данными, в результате чего в уравнении появляются слишком большие коэффициенты разных знаков. Количественные характеристики зависимости определяются из корреляционной матрицы, которая также может служить матрицей системы уравнений. Для уменьшения коэффициентов корреляции традиционно используется Ridge (гребневая) регрессия, которая предполагает увеличение диагональных членов матрицы на одно и то же положительное число. В результате число обусловленности матрицы уменьшается и уравнение регрессии становится более устойчивым: небольшое изменение входных данных приводит к небольшому изменению решения. Число, на которое увеличиваются диагональные члены матрицы, называется штрафом, накладываемым в гребневой регрессии на все коэффициенты регрессии. В предлагаемом методе штрафы, причем разные, налагаются только на те коэффициенты, которые соответствуют данным с высокой корреляцией. Это приводит к повышению устойчивости уравнения вследствие уменьшения значений коэффициентов, соответствующих коррелированным данным. Выбор элементов, подлежащих увеличению, основывается на анализе корреляционной матрицы исходного набора данных с помощью разложения ее на диагональные матрицы методом квадратных корней. Кроме повышения устойчивости с помощью предлагаемого метода может быть достигнуто снижение размерности регрессионной модели – уменьшение количества членов соответствующего уравнения, для чего обычно используются алгоритмы LASSO и LARS. Эффективность метода проверяется на известном наборе данных, причем выполняется сравнение не только с Ridge-регрессией, но и с результатами известных алгоритмов уменьшения размерности.

Линейная регрессия; нормальные уравнения; корреляционная матрица; штрафы на отдельные коэффициенты уравнения.

V.N. Lutay, N.S. Khusainov

INCREASING THE STABILITY OF LINEAR REGRESSION

The paper considers a method for increasing the stability of a linear regression equation, the coefficients of which are the solution of a system of normal equations. The instability of the equation is determined by the presence of a linear relationship between the experimental data; the quantitative characteristics of the dependence are determined from the correlation matrix, which can also serve as a matrix of the system of equations. To reduce the correlation coefficients, Ridge (ridge) regression is traditionally used, which involves an increase in the diagonal members of the matrix by the same positive number. As a result, the matrix condition number decreases and the regression equation becomes more stable: a small change in the input results in a small change in the solution. The number by which the diagonal terms of the matrix increase is called the penalty imposed in ridge regression on all regression coefficients. In the proposed method, penalties, and different ones, are imposed only on those coefficients that correspond to data with high correlation. This leads to an increase in the stability of the equation due to a decrease in the values of the coefficients corresponding to correlated data. The choice of elements to be increased is based on the analysis of the correlation matrix of the original data set by decomposing it into diagonal matrices using the square root method. In addition to increasing the stability using the proposed method, a reduction in the dimension of the regression model can be achieved - a decrease in the number of terms of the corresponding equation, for which the LASSO and LARS algorithms are usually used. The effectiveness of the method is tested on a known data set, and a comparison is made not only with Ridge regression, but also with the results of known dimensionality reduction algorithms.

Linear regression; normal equations; correlation matrix; penalties on individual equation coefficients.

Введение. Регрессионные модели широко применяются для анализа и предсказания данных. В работах [1–6] излагаются основы теории, в [8–13] ее применение. Алгоритмы анализа и построения регрессионных моделей являются частью алгоритмов машинного обучения.

Линейная регрессионная модель отражает связь между значениями независимых переменных (предикторами) и объясняющими переменными (откликами):

$$y_i = \beta_0 + \sum_{j=1}^{m} \beta_j x_{ij}\,,\ i = \overline{1,n}\,,\ \ n \geq m,$$

где x_{ij} – наблюдаемое значение предиктора x_j, y_i – значение отклика, β_j – коэффициент регрессионного уравнения. После центрирования наблюдаемых значений [4] линейную регрессионную модель удобно записать в матрично-векторной форме (без β_0)

$$W\bar{b} = \bar{v},$$

где $W = [w_{ij}]_{nxm}$ – прямоугольная матрица значений предикторов, $\bar{b} = [b_j]_m$ и $\bar{v} = [v_i]_n$– векторы коэффициентов модели и значений откликов соответственно. Традиционно для получения значений b_j используется метод наименьших квадратов, который сводится к решению системы, называемой нормальными уравнениями, с положительно определенной симметричной матрицей.

$$(W^t W)\hat{b} = W^t \bar{v},$$

где t – символ транспонирования. После определения элементов $\hat{b}$ и обратного перехода к наблюдаемым значениям получим линейную МНК-модель:

$$\hat{y}_i = \hat{\beta}_0 + \sum_{j=1}^{m} \hat{\beta}_j x_{ij}\,, i = \overline{1,n}, n \geq m.$$

Эффективность модели измеряется по сумме квадратов отклонений полученного решения от значений отклика:

$$SSE = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2.$$

Если после центрирования выполнить нормирование наблюдаемых переменных, то $W^t W$становится матрицей парных корреляций предикторов. Диагональные элементы матрицы равны 1, все значения заключены между -1 и 1. По значениям недиагональных членов матрицы можно сделать заключение о взаимной корреляции (коллинеарности) предикторов. Чем больше корреляция между i-м и j-м предикторами, тем ближе значение недиагонального элемента с индексом ij к 1 и тем больше число обусловленности матрицы и дисперсия значений откликов. Если такой элемент матрицы равен 1, то соответствующие предикторы линейно зависимы и матрица сингулярна. Корреляционная матрица зачастую используется в качестве матрицы нормальных уравнений.

Традиционно для анализа и построения регрессионной модели используются алгоритмы Ridge, LASSO, LARS. В Ridge-модели, используемой в случае мультиколлинеарных наборов данных, к матрице W^tWдобавляется диагональная матрица λI

$$(W^t W + \lambda I)\bar{b}_\lambda = W^t \bar{v}, \tag{1}$$

где $\bar{b}_\lambda = [b(\lambda)_j]_m$, λ – положительное число, называемое параметром.

Добавление параметра ко всем диагональным членам положительно определенной симметричной матрицы приводит к уменьшению ее числа обусловленности. Значения λ называют штрафом, налагаемым на коэффициенты модели:

$$P(\lambda) = \lambda \sum_{j=1}^{m} b_j^2.$$

Ridge- модель в этом случае приобретает следующую форму:

$$y(\lambda)_i = \beta(\lambda)_0 + \sum_{j=1}^{m} \beta(\lambda)_j \, x_{ij}. \tag{2}$$

При увеличении λ дисперсия уменьшается, но возрастает смещение коэффициентов регрессии β(λ) от значений МНК, так как при λ→∞ значения $b_\lambda \to 0$. Ridge-коэффициенты можно рассматривать как линейные комбинации МНК-оценок.

LASSO [14, 15] заключается во введении ограничения на норму вектора коэффициентов модели, что приводит к обращению в 0 некоторых из них. Алгоритм LARS (метод наименьших углов) [16] может уменьшать размерность модели, анализируя вклад предикторов в отклик аналогично методу пошаговой регрессии. Последний в [17] рассматривается как наилучший способ построения линейной модели.

Далее рассматривается метод понижения числа обусловленности матрицы, отличный от Ridge-модели тем, что увеличиваются не все диагональным члены матрицы, а только те, которые являются причиной большой коллинеарности модели. При этом добавляемые параметры для разных членов могут быть различными. Такой способ регуляризации будем называть выборочным (SR).

Выбор диагональных элементов. Обозначим *A* матрицу W^tW и запишем результат треугольной декомпозиции ее методом Холецкого, который является одним из популярных методов решения систем ЛАУ с симметричной матрицей:

$$A = LL^t,$$

где L – верхняя треугольная матрица.

Число обусловленности матрицы *A* можно уменьшить, если уменьшить число обусловленности ее множителей, так как:

$$cond(A) \le cond(L)^2.$$

В [18] показано, что нижняя граница числа обусловленности треугольной невырожденной матрицы имеет следующий вид:

$$cond(L) \ge \frac{\max\limits_{ij} |l_{ij}|}{\min\limits_{ij} |l_{ij}|}.$$

Таким образом, для уменьшения числа обусловленности матрицы A следует увеличивать минимальный диагональный элемент L и не увеличивать максимальный. Ниже приведены формулы для вычисления диагональных элементов треугольной матрицы *L* [19].

$$l_{ii} = \sqrt{a_{ii} - \sum_{k=1}^{i-1} l_{ik}^2}\,, i = 2, n. \tag{3}$$

Для получения коэффициентов системы уравнений решаются две системы с треугольными матрицами

$$L\bar{z} = W^t\bar{v}, L^t\hat{b} = \bar{z}. \qquad (4)$$

В [20] показано, что увеличение i-го диагонального элемента треугольной матрицы в процессе треугольного разложения равносильно тому, что разложению подвергается матрица A с увеличенным a_{ii} без изменения остальных диагональных элементов. Из этого следует, что если в процессе факторизации произошло увеличение нескольких диагональных элементов треугольной матрицы, то такое разложение является точным разложением следующей полной матрицы

$$LL^t = A + S,$$

где S – матрица, ненулевые элементы которой расположены на главной диагонали и их индексы совпадают с индексами тех диагональных элементов треугольной матрицы, которые увеличиваются в процессе разложения. Верно и обратное: из (3) следует, что при увеличении a_{ii} на δa_{ii} значение l_{ii} увеличивается на δl_{ii} согласно следующему соотношению (все величины неотрицательные):

$$\delta l_{ii}^2 + 2\delta l_{ii} = \delta a_{ii}$$

Из вышесказанного вытекает, что увеличением a_{ii} может уменьшить число обусловленности A.

Рассмотрим вопрос о выборе подходящего a_{ii}. С этой целью обратимся к результату разложения матрицы по Холецкому. В качестве примера ниже приведены соответствующие матрицы для n=3, а в (5) – выражения для диагональных элементов треугольной матрицы.

$$\begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix} \cdot \begin{bmatrix} l_{11} & l_{21} & l_{31} \\ 0 & l_{22} & l_{32} \\ 0 & 0 & l_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix}$$

$$l_{11} = 1,$$
$$l_{22}^2 = a_{22} - a_{12}^2, \qquad (5)$$
$$l_{33}^2 = a_{33} - a_{13}^2 - (a_{23} - a_{12}a_{13})^2 / l_{22}^2.$$

Из приведенных выражений следует, что если диагональные элементы A равны 1 (как в корреляционной матрице), то величина l_{ii} зависит от элементов матрицы A в i-ом столбце (a_{12} и a_{13} соответственно): если в i-ом столбце корреляционной матрицы находится большое число, то l_{ii} мало. Увеличение этого члена приводит согласно (4) к уменьшению соответствующего коэффициента регрессионного уравнения и, следовательно, к уменьшению влияния соответствующего предиктора на отклик.

После увеличения одного или нескольких диагональных элементов корреляционной матрицы, выбранных по соответствующей треугольной матрице, по аналогии с (1) можно записать

$$(W^tW + S)\bar{b}_s = W^t\bar{v}.$$

Регрессионная модель в этом случае выглядит так

$$y_i(s) = \sum_{j=0}^{m} \beta_j(s) x_{ij}, \qquad (6)$$
$$i = \overline{1, n}$$

Положим, что количество ненулевых диагональных членов в S равно k. Тогда штраф налагается не на все параметры модели, а только на k из них:

$$P(s) = \sum_{j=1}^{k} s_{jj} b_j^2 .$$

Вычислительные эксперименты. Для анализа эффективности предлагаемого метода были использованы данные Хальда, подробно проанализированные в [8]. В этом наборе данных число наблюдений 13, количество предикторов 4. Корреляционная матрица имеет следующий вид (все вычисления и построение графиков выполнены на языке Python):

$$\begin{bmatrix} 0.999 & 0.228 & -0.824 & -0.245 \\ 0.228 & 1.0 & -0.139 & -0.972 \\ -0.824 & -0.139 & 0.999 & 0.029 \\ -0.245 & -0.972 & 0.029 & 1.0 \end{bmatrix}.$$

Из вида матрицы следует, что взаимная корреляция велика между 1-м и 3-м предикторами и между 2-м и 4-м. Значения диагональных элементов треугольной матрицы соответствуют этому:

$$l_{11} = 0.999; l_{22} = 0.973; l_{33} = 0.564; l_{44} = 0.0.0594.$$

Таким образом, для снижения числа обусловленности матрицы следует увеличивать ее 3-й и 4-й диагональные элементы.

При выполнении численных расчетов следует иметь в виду, что при увеличении значения $l_{ii}, i > 1$ посредством увеличения a_{ii} недиагональные члены l_{ji}уменьшаются, а диагональные $l_{jj}, j > i$ увеличиваются. При этом элемент треугольной матрицы l_{11} остается равным 1. Поэтому желательно, чтобы диагональные элементы треугольной матрицы с возрастанием i уменьшались так, чтобы минимальный член был последним. Это можно сделать посредством перестановки столбцов предикторов. В нашем случае это условие выполнено.

В табл. 1 приведены характеристики алгоритма SR при увеличении параметра λ и соответствующих диагональных элементов матрицы в сравнении с Ridge-регрессией.

Таблица 1

Сравнительные характеристики выборочной регуляризации и Ridge

№	*cond(W^tW)*	λ	δa_{33}	δa_{44}	Смещение оценки	
					R	SR
1	1400	0	0	0	0	0
2	400	0.004	0.005	0.007	0.26	0.15
3	140	0.015	0.01	0.03	0.33	0.21
4	60	0.038	0.05	0.07	0.35	0.21
5	18	0.12	0.1	0.3	0.38	0.23
6	8	0.35	0.5	0.8	0.42	0.23
7	7	0.45	0.7	1	0.44	0.23
8	7	0.55	0.8	1.1	0.45	0.23

Второй столбец таблицы – числа обусловленности матрицы, третий – значения параметра λ, соответствующего этим числам, Далее идут столбцы значений, добавляемых к 3-му и 4-му диагональным элементам исходной корреляционной матрицы, а в столбцах R и SR приведены значения смещения математического ожидания для Ridge-регрессии и предлагаемого метода.

На рис. 1 приведены графики изменения $\beta_i(s)$ из (6) (сплошные линии) и $\beta_i(\lambda)$ из (2) (пунктирные) в зависимости от уменьшения числа обусловленности (цифры на оси x соответствуют первому столбцу табл. 1). Рисунок показывает, что $\beta_3(s)$ и $\beta_4(s)$ имеют меньшие абсолютные значения, чем $\beta_3(\lambda)$ и $\beta_4(\lambda)$. Значения β_0, не приведенные на рис. 1, снижаются с уменьшением числа обусловленности с 55 до 51 для SR и повышаются с 75 до 89 для Ridge-регрессии.

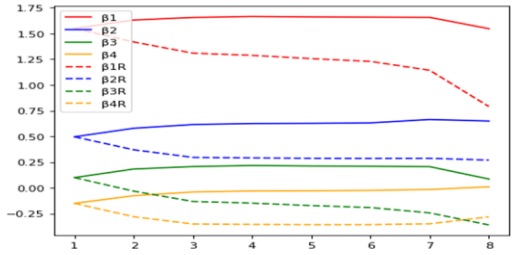

Рис. 1. Коэффициенты регрессии для Ridge и SR.

Оптимальное значение λ, выбранного в [17] для данных Хальда, равно 0.013. В табл. 1 этому значению соответствует точка 3; значения коэффициентов регрессии для Ridge и SR сведены в табл. 2. Из последнего столбца следует, что SR уменьшает β_4 в соответствии с тем обстоятельством, что четвертый предиктор наиболее коррелирован.

Таблица 2

Коэффициенты регрессии для λ= 0.013.

	$\boldsymbol{\beta_0}$	$\boldsymbol{\beta_1}$	$\boldsymbol{\beta_2}$	$\boldsymbol{\beta_3}$	$\boldsymbol{\beta_4}$
SR	50.746	1.674	0.629	0.227	-0.0268
Ridge	83.437	1.308	0.297	-0.133	-0.351

На рис. 2 показаны зависимости значений SSE при уменьшении числа обусловленности матрицы. Значения SSE для Ridge-регрессии резко возрастают при увеличении параметра регуляризации.

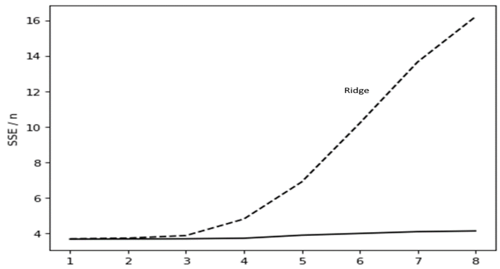

Рис. 2. Значения SSE для Ridge регрессии(пунктирная линия) и для SR (сплошная)

На рис. 3 приведены графики, отражающие результаты сравнения SR со значениями, полученными при использовании функции linear_model.LassoLars(), которая реализует совместно алгоритмы Lasso и LARS. Далее такое объединение сокращенно называется Lars. Единицами оси x на этом и на следующим рисунке выбраны приведенные в таблице 3 номера, каждому из которых соответствует значения параметра α в алгоритме Lars и значения δa_{33} и δa_{44} (так как их начальным значением в табл. 3 является последнее значение табл. 1, то ось *x* на рис. 3 и 4 можно рассматривать как продолжение той же оси на рис. 1 и 2.)

Таблица 3

Соответствие параметра Lars и значений изменяемых диагональных элементов членов матрицы

№	8	9	10	11	12	13	14	15
α (LARS)	0.01	0.015	0.018	0.02	0.025	0.03	0.035	0.05
$\boldsymbol{\delta a_{33}(\delta a_{44})}$	1.1	1.2	1.5	1.8	2.2	3.0	3.3	4.0

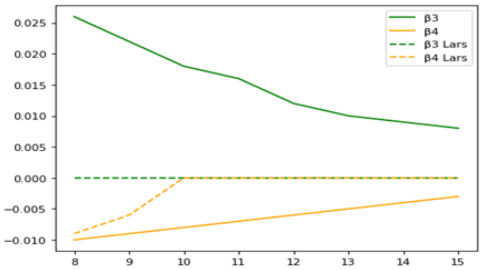

Рис. 3. Коэффициенты регрессии для Lars и SR

Значения β_3 и β_4 для LARS, начиная с точки 10, равны 0. В SR эти коэффициенты малы, но 0 не равны. С возрастанием δa_{33} и δa_{44} их абсолютные значения уменьшаются; в точке 14 по сравнению со значениями β_1 и β_2 становятся незначительными и могут не учитываться в регрессионном уравнении.

На рис. 4 показаны изменения SSE для Lars и SR. Из него следует, что SSE в точке 14, в которой для SR β_3 и β_4 были приняты равными 0, меньше, чем у Lars в точке 10.

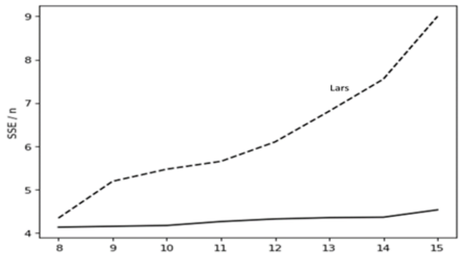

Рис. 4. Изменения SSE для LARS и SR

Заключение. В работе получены следующие результаты:

♦ показано, что увеличение нескольких диагональных членов корреляционной матрицы или матрицы нормальных уравнений приводит к увеличению устойчивости регрессионного уравнения так же, как и в Ridge-регрессии при увеличении всех диагональных элементов;

♦ в отличие от Ridge увеличение диагональных элементов приводит к уменьшению значений коэффициентов регрессии, соответствующих коллинеарным предикторам;

♦ эффективность метода в плане уменьшения размерности модели сравнима с известным алгоритмами Lasso и LARS; при этом метод существенно проще и показывает на приведенных экспериментальных данных лучший результат по среднеквадратическому отклонению.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Ohlsson H., Ljung L.* Identification of switched linear regression mod-els using sum-of-norms regularization // Automatica. – 2013. – Vol. 149 (4). – P. 1045-50.
2. *Hahn P.R., Carvalho C.M., Puelz D., He J.* Regularization and confounding in linear regression for treatment effect estimation // Bayesian Analysis. – 2018. – Vol. 13 (1). – P. 163-82.
3. *Cui Q., Xu Y., Zhang Z., Chan V.* Max-linear regression models with regu-larization // Journal of Econometrics. – 2021. – Vol. 222 (1). – P. 579-600.
4. *Yan X.* Linear Regression Analysis: Theory and Computing. – World Scientific Publishing Company, 2009.
5. *Альберт А*. Регрессия, псевдоинверсия и рекуррентное оценивание: пер. с англ. – М.: Наука. Главная редакция физико-математической литературы, 1977. – 224 с.
6. *Себер Дж.* Линейный регрессионный анализ: пер, с англ. / под ред. М.Б. Малютова. – М.: Мир.
7. *Носко В.П.* Эконометрика: учебник. – М.: ИД Дело РАНХиГС, 2011. – 672 с.
8. *Успенский А.Б., Федоров В.В.* Вычислительные аспекты метода наименьших квадратов при анализе и планировании регрессионных экспериментов. – М.: Изд-во МГУ, 1975.
9. *Горидько Н.П., Нижегородцев Р.М.* Современный экономический рост: теория и регрессионный анализ: монография. – М.: Инфра-М, 2017. – 444 с.
10. *Есаулов И.Г.* Регрессионный анализ данных в пакете Mathcad: учеб. пособие. – СПб.: Лань П, 2016. – 224 с.
11. *Карлберг К.* Регрессионный анализ в Microsoft Excel. – М.: Диалектика, 2019. – 400 с.
12. *Соколов Г.А.* Введение в регрессионный анализ и планирование регрессионных экспериментов в экономике: учеб. пособие. – М.: Инфра-М, 2016. – 352 с.
13. Факторный анализ преступности: корреляционный и регрессионный методы: монография / под ред. С.М. Иншакова. – М.: Юнити, 2014. – 127 с.
14. *Trevor H.,Tibshirani, R,Wainwright,M.* Statistical Learning with Sparsity: The Lasso and Generalizations. – Chapman&Hall, 2015.
15. *Мелкумова Л.Э., Шатских С.Я.* Сравнение методов Ридж-регрессии и LASSO в задачах обработки данных // Сб. трудов III международной конференции и молодежной школы. Самарский национальный исследовательский университет имени академика С.П. Королева. – 2017. – С. 1755-62.
16. *Efron B., Hastie T., Johnstone J., Tibshirani R.* Least Angle Regression // The Annals of Statistics. – 2004. – Vol. 32. – P. 407-99.
17. *Дрейпер Н., Смит Г.* Прикладной регрессионный анализ. – 2-е изд.: пер. с англ. – М.: Финансы и статистика, 1986.
18. *Лоусон Ч., Хенсон Р.* Численное решение задач метода наименьших квадратов: пер, с англ. – М.: Наука. Гл. ред. физ.-мат. лит., 1986.
19. *Воеводин В.В., Кузнецов Ю.А.* Матрицы и вычисления. – М.: Наука. Главная редакция физико-математической литературы, 1984.
20. *Лутай В.Н.* Повышение устойчивости треугольного разложения плохо обусловленных матриц // СибЖВМ. – 2019. – № 4. – Т. 22. – С 465-73.

REFERENCES

1. *Ohlsson H., Ljung L.* Identification of switched linear regression mod-els using sum-of-norms regularization, *Automatica*, 2013, Vol. 149 (4), pp. 1045-50.
2. *Hahn P.R., Carvalho C.M., Puelz D., He J.* Regularization and confounding in linear regression for treatment effect estimation, *Bayesian Analysis*, 2018, Vol. 13 (1), pp. 163-82.
3. *Cui Q., Xu Y., Zhang Z., Chan V.* Max-linear regression models with regu-larization, *Journal of Econometrics*, 2021, Vol. 222 (1), pp. 579-600.
4. *Yan X.* Linear Regression Analysis: Theory and Computing. World Scientific Publishing Company, 2009.
5. *Al'bert A.* Regressiya, psevdoinversiya i rekurrentnoe otsenivanie [Regression, pseudo–inversion and recurrent estimation]: trans. from engl. Moscow: Nauka. Glavnaya redaktsiya fiziko-matematicheskoy literatury, 1977, 224 p.
6. *Seber Dzh.* Lineynyy regressionnyy analiz [Linear regression analysis]: trans. from engl., ed. by M.B. Malyutova. Moscow: Mir.

7. *Nosko V.P.* Ekonometrika: uchebnik [Econometrics: textbook]. Moscow: ID Delo RANKHiGS, 2011, 672 p.
8. *Uspenskiy A.B., Fedorov V.V.* Vychislitel'nye aspekty metoda naimen'shikh kvadratov pri analize i planirovanii regressionnykh eksperimentov [Computational aspects of the least squares method in the analysis and planning of regression experiments]. Moscow: Izd-vo MGU, 1975.
9. *Gorid'ko N.P., Nizhegorodtsev R.M.* Sovremennyy ekonomicheskiy rost: teoriya i regressionnyy analiz: monografiya [Modern economic growth: theory and regression analysis: monograph]. Moscow: Infra-M, 2017, 444 p.
10. *Esaulov I.G.* Regressionnyy analiz dannykh v pakete Mathcad: ucheb. posobie [Regression analysis of data in the Mathcad package: tutorial]. Saint Petersburg: Lan' P, 2016, 224 p.
11. *Karlberg K.* Regressionnyy analiz v Microsoft Excel [Regression analysis in Microsoft Excel]. Moscow: Dialektika, 2019, 400 p.
12. *Sokolov G.A.* Vvedenie v regressionnyy analiz i planirovanie regressionnykh eksperimentov v ekonomike: ucheb. posobie [Introduction to regression analysis and planning of regression experiments in economics: textbook]. Moscow: Infra-M, 2016, 352 p.
13. Faktornyy analiz prestupnosti: korrelyatsionnyy i regressionnyy metody: monografiya [Factor analysis of crime: correlation and regression methods: monograph], ed. by S.M. Inshakova. Moscow: Yuniti, 2014, 127 p.
14. *Trevor H.,Tibshirani, R,Wainwright,M.* Statistical Learning with Sparsity: The Lasso and Generalizations. Chapman&Hall, 2015.
15. *Melkumova L.E., Shatskikh S.Ya.* Sravnenie metodov Ridzh-regressii i LASSO v zadachakh obrabotki dannykh [Comparison of Ridge regression and LASSO methods in data processing problems], Cb. trudov III mezhdunarodnoy konferentsii i molodezhnoy shkoly [Proceedings of the III International Conference and Youth School]. Samara National Research University named after Academician S.P. Korolev, 2017, pp. 1755-62.
16. *Efron B., Hastie T., Johnstone J., Tibshirani R.* Least Angle Regression, T*he Annals of Statistics,* 2004, Vol. 32, pp. 407-99.
17. *Dreyper N., Smit G.* Prikladnoy regressionnyy analiz [Applied Regression analysis]. 2nd ed.: trans. from engl. Moscow: Finansy i statistika, 1986.
18. *Louson Ch., Khenson R.* Chislennoe reshenie zadach metoda naimen'shikh kvadratov [Numerical solution of problems of the least squares method]: trans. from engl. Moscow: Nauka. Gl. red. fiz.-mat. lit., 1986.
19. *Voevodin V.V., Kuznetsov Yu.A.* Matritsy i vychisleniya [Matrices and calculations]. Moscow: Nauka. Glavnaya redaktsiya fiziko-matematicheskoy literatury, 1984.
20. *Lutay V.N.* Povyshenie ustoychivosti treugol'nogo razlozheniya plokho obuslovlennykh matrits [Increasing the stability of the triangular decomposition of poorly conditioned matrices], *SibZhVM* [Siberian Journal of Computational Mathematics], 2019, No. 4, Vol. 22, pp. 465-73.

Статью рекомендовал к опубликованию д.т.н., профессор В.М. Курейчик.

Лутай Владимир Николаевич – Южный федеральный университет; e-mail: vnlutay@sfedu.ru; г. Таганрог, Россия; тел.: 89282276002; кафедра математического обеспечения и применения ЭВМ; к.т.н.; доцент.

Хусаинов Наиль Шявкатович – e-mail: khusainov@sfedu.ru; тел.: 89034613347; кафедра математического обеспечения и применения ЭВМ; зав.кафедрой; к.т.н.; доцент.

Lutay Vladimir Nikolaevich – Southern Federal University; e-mail: vnlutay@sfedu.ru; Taganrog, Russia; phone: +79282276002; the department of software engineering; cand. of eng. sc.; associated professor.

Khusainov Nail' Shyavkatovich – e-mail: khusainov@sfedu.ru; phone: +79034613347; the department of computer aided design; head of department; cand. of eng. sc.; associated professor.