УДК 519.237.5

А. М. Михеенко, Д. С. Савич

ПРИМЕНЕНИЕ ВЕКТОРА ШЕПЛИ В РЕГРЕССИОННОМ АНАЛИЗЕ

Рассматривается применение модели кооперативной теории игр, вектора Шепли, для оценки коэффициентов регрессии и вклада предикторов в общую объясняемость модели. Систематизируется необходиый теоретический аппарат и приводится построение оценок Шепли, основываясь на коэффициенте детерминации.

The application of the cooperative game theory model, Shapley value, for the estimation of the regression coefficients and the contribution of predictors to the general explainability of the model is considered. The necessary theoretical apparatus is systematized and the construction of Shapley estimates based on the coefficient of determination is given.

Ключевые слова: вектор Шепли, мультиколлинеарность, линейная регрессия, кооперативная теория игр, коэффициент детерминации.

Keywords: Shapley value, multicollinearity, linear regression, cooperative game theory, coefficent of determination.

Введение

Впервые термин «регрессия» появился в конце XIX в. в работе Френсиса Гальтона, а впоследствии сформировался так называемый регрессионный анализ. На данный момент он является одним из наиболее распространенных методов обработки наблюдений и исследования взаимосвязей между переменными в различных сферах.

Одной из задач регрессионного анализа является нахождение неизвестных коэффициентов регрессии. Для решения данной задачи разработано огромное количество различных методов, но наиболее часто используемый из них — это метод наименьших квадратов [14].

Однако при определенных условиях оценки этого метода становятся нестабильными. Этим ограничивающим условием является мультиколлинеарность, или достаточно сильная линейная зависимость между предикторами [13]. Для уменьшения деструктивного воздействия такой зависимости используют различные методы [10]. Об одном из них речь и пойдет далее.

Данный метод основывается на применении аппарата кооперативной теории игр [1; 5]. А точнее, на одной из его конструкций, называемой вектором Шепли и придуманной нобелевским лауреатом Ллойдом Шепли в 1953 г. [15]. Этот вектор сопоставляет каждому «игроку» его выигрыш, усредненный по всевозможным объединениям таких же игроков.

Вестник Балтийского федерального университета им. И. Канта.
Сер.: Физико-математические и технические науки. 2020. № 2. С. 84 – 94.

Возникает вопрос — в чем преимущества данного метода? Важнейшим моментом является то, что оценки, получаемые с помощью вектора Шепли, не подвергаются воздействию мультиколлинеарности и имеют ряд замечательных свойств.

Соответственно, данная работа посвящена анализу применения вектора Шепли в регрессионном анализе. Аналогичная ей уже проводилась С. Липовецким и М. Конклином [8; 12]. В данной статье мы постараемся собрать всю необходимую информацию, систематизировать ее и обобщить.

Классический регрессионный анализ

В настоящей работе ограничимся рассмотрением лишь линейных моделей. Итак, модель $y=\beta_0+\beta_1x_1+\beta_2x_2+\ldots++\beta_nx_n+\varepsilon$ называют моделью множественной линейной регрессии. Прилагательное «линейная» используется для обозначения линейности по параметрам β_j, $j=\overline{1,n}$, а не потому что y есть линейная функция от x [2].

Рассмотрим выборочную регрессионную модель. То есть предположим, что проводится N наблюдений за выявлением значения отклика y_i при фиксировании каких-либо значений x_j. Очевидно, что при таких наблюдениях будет присутствовать случайная ошибка ε_i, тогда первоначальную модель можно переписать в следующем виде:

$$y_i=\beta_1x_{i1}+\beta_2x_{i2}+\ldots++\beta_nx_{in}+\varepsilon_i=\sum_{j=1}^{n}\beta_jx_{ij}+\varepsilon_i,$$

где $i=\overline{1,N}$ — номер наблюдения.

Для упрощения вычислений и записи будем пользоваться матричными обозначениями:

$$Y=X\beta+\varepsilon,$$

$$Y=\begin{pmatrix}y_1\\y_2\\\vdots\\y_N\end{pmatrix},X=\begin{pmatrix}x_{11}&x_{12}&\cdots&x_{1n}\\x_{21}&x_{22}&\cdots&x_{2n}\\\vdots&\vdots&\cdots&\vdots\\x_{N1}&x_{N2}&\cdots&x_{Nn}\end{pmatrix},\beta=\begin{pmatrix}\beta_1\\\beta_2\\\vdots\\\beta_n\end{pmatrix},\varepsilon=\begin{pmatrix}\varepsilon_1\\\varepsilon_2\\\vdots\\\varepsilon_N\end{pmatrix},$$

$$Y=\begin{pmatrix}y_1\\y_2\\\vdots\\y_N\end{pmatrix},X=\begin{pmatrix}x_{11}&x_{21}&\cdots&x_{n1}\\x_{12}&x_{22}&\cdots&x_{n2}\\\vdots&\vdots&\cdots&\vdots\\x_{1N}&x_{2N}&\cdots&x_{nN}\end{pmatrix},\beta=\begin{pmatrix}\beta_1\\\beta_2\\\vdots\\\beta_n\end{pmatrix},\ \mathcal{E}=\begin{pmatrix}\varepsilon_1\\\varepsilon_2\\\vdots\\\varepsilon_N\end{pmatrix}.$$

Предположения:

1) $\varepsilon_i, i=\overline{1,N}$ — случайная величина, причем ее математическое ожидание равно нулю, а ее значения некоррелированы и имеют одинаковые дисперсии σ^2, то есть

$$E(\varepsilon_i)=0; Var(\varepsilon_i)=\sigma^2; \operatorname{cov}(\varepsilon_i,\varepsilon_j)=0, \text{при}\,(i\neq j)\wedge\left(i,j=\overline{1,N}\right);$$

2) матрица X детерминирована, то есть x_{ji} — не являются случайными величинами, и ее ранг равен числу коэффициентов в модели: $rank(X)=n, N\geq n$;

3) на значения β_j в модели не налагается каких-либо ограничений.

Так как результаты наблюдений есть величины случайные, то значения β_j получить не удастся, но на основе эмпирических наблюдений можно получить для них статистические оценки. Такая модель, где значения параметров заменены их оценками, имеет вид $\widehat{Y}=Xb$, где b — вектор оценок параметров; $\widehat{Y}$ — предсказанное значение отклика.

$$b=\begin{pmatrix} b_1\\ b_2\\ \vdots\\ b_n\end{pmatrix}, \widehat{Y}=\begin{pmatrix} \widehat{y_1}\\ \widehat{y_2}\\ \vdots\\ \widehat{y_N}\end{pmatrix}.$$

Под действием случайных возмущений значение $\widehat{y_i}$ будет отличаться от результата измерения y_i. Тогда введем разности, которые будем называть остатками или невязками: $e_i=y_i-\widehat{y_i}, i=\overline{1,N}$.

Для нахождения коэффициентов регрессии естественно потребовать минимизацию остатков так называемой функции риска [4].

Наиболее часто используемая функция:

$$S^2=\sum_{i=1}^{N}\left(y_i-\widehat{y_i}\right)^2.$$

Для компенсации различия в знаках в ней остатки возводятся в квадрат. Запишем ее в матричном виде:

$$S^2=e^te=\left(Y-\widehat{Y}\right)^t\left(Y-\widehat{Y}\right)=\left(Y-Xb\right)^t\left(Y-Xb\right)=$$
$$=Y^tY-Y^tXb-\left(Xb\right)^tY+b^tX^tXb=Y^tY-2Y^tXb+b^tX^tXb\to\min,$$

где e^t означает транспонированную матрицу e. Метод, оценивающий регрессионные коэффициенты и заключающийся в минимизации величины S^2, называется методом наименьших квадратов, или МНК.

Итак, в силу предположения 3 на возможные значения оценок не налагается никаких ограничений. Тогда будем искать минимум, приравняв к нулю частные производные по b: $\dfrac{\partial S^2}{\partial b}=-2X^tY+2X^tXb=0$.

Получим систему нормальных уравнений в матричном виде:

$$X^t Xb = X^t Y.$$

Считая, что предположение 2 выполнено (а точнее, $rank(X) = n$, что необходимо для существования обратной матрицы) и умножив слева обе части на обратную к $X^t X$ матрице, получим оценку метода наименьших квадратов:

$$b = \left(X^t X\right)^{-1} X^t Y.$$

Введем некоторые обозначения, которые будут использованы позже: матрицу $X^t X = C$ принято называть информационной матрицей, или матрицей плана, а $\left(X^t X\right)^{-1} = C^{-1}$ — матрицей ошибок, или матрицей дисперсий-ковариаций; $X^t Y = r$. Окончательно получим оценку МНК:

$$b = C^{-1} r.$$

Оценка МНК является несмещенной; эффективной в классе линейных несмещенных оценок; состоятельной [3].

Перейдем к критериям качества модели, а точнее к одному из них, называемым коэффициентом детерминации, или R^2:

$$R^2 = 1 - \frac{\sum_{i=1}^N \left(y_i - \widehat{y}_i\right)^2}{\sum_{i=1}^N \left(y_i - \overline{y}\right)^2} = 1 - \frac{\sum_{i=1}^N e_i^2}{\sum_{i=1}^N \left(y_i - \overline{y}\right)^2}.$$

Знаменатель в формуле принято называть общей суммой квадратов, а числитель — суммой квадратов остатков. Можно переписать формулу для R^2 в следующем виде, где в числителе будет присутствовать так называемая объясненная сумма квадратов:

$$R^2 = \frac{\sum_{i=1}^N \left(\widehat{y}_i - \overline{y}\right)^2}{\sum_{i=1}^N \left(y_i - \overline{y}\right)^2}.$$

В зависимости от характера переменных x_{ij} коэффициент детерминации может характеризоваться по-разному. В общем случае он показывает, какая доля вариации отклика учтена в модели и объясняется действиями предикторов. Однако, если все x_{ij} есть величины детерминированные, то R^2 показывает, насколько данная регрессионная модель лучше регрессии к среднему.

Геометрически данный коэффициент есть отношение квадрата длины катета к квадрату гипотенузы. Другими словами, R^2 равен квадрату косинуса угла между Y и $\widehat{Y}$. Рассмотрим его свойства:

1) $0 \le R^2 \le 1$;

2) чем ближе R^2 к единице, тем сильнее зависимость Y и X;

3) $R^2 = 1$ лишь когда отклик и предикторы связаны функциональной зависимостью;

4) если $R^2 = 0$, то выбранная модель регрессии качественно не превосходит модель $\hat{y} = \bar{y}$.

Здесь и далее будем предполагать, что все x_j центрированы и нормированы на единичную длину.

Мультиколлинеарностью будем называть сопряженность (подразумевая детерминированность независимых переменных) предикторов. Различают сильную мультиколлинеарность и слабую, обычно ее называют просто мультиколлинеарностью.

Сильная мультиколлинеарность выражается в функциональной линейной зависимости между предикторами, а слабая характеризуется сильной корреляцией между объясняющими переменными.

Далее мы будем опускать слово «слабая», подразумевая его. Мультиколлинеарность характеризуется сильной корреляцией между объясняющими переменными.

Отрицательные эффекты мультиколлинеарности следующие:

1) хотя матрица $X^t X = C$ и будет невырожденной, ее определитель будет достаточно мал, что влечет за собой большие значения обратной матрицы и, как следствие, большие дисперсии σ^2 для оценок b_j;

2) неустойчивость оценок, то есть добавление или исключение небольшого количества информации, приводит к сильному изменению коэффициентов и уменьшению точности предсказания по модели;

3) возможно получение не только неправильных коэффициентов с точки зрения теории, но и противоположного знака;

4) уменьшение t-статистик коэффициентов b_j и оценка их значимости теряет смысл.

Вектор Шепли

Для построения вектора Шепли необходимо знать определенный минимум информации. Начнем с определения характеристической функции в терминах кооперативной теории игр.

Определение. Характеристическая функция (х.ф.) над множеством I — отображение $v : 2^I \to \mathbb{R}$. Будем предполагать, что $I = \{1, ..., n\}$ конечно, а его элементы называть игроками.

Итак, будем считать, что характеристическая функция ставит в соответствие каждому подмножеству некоторого множества игроков вещественное число. Любое такое подмножество будем называть коалицией.

Ограничимся рассмотрением характеристических функций, обладающих некоторыми логически очевидными свойствами.

Свойства:

1) персональность: $\nu(\varnothing)=0$;
2) супераддитивность: $\forall K,L\subset I: K\cap L=\varnothing \ \ \nu(K)+\nu(L)\le\nu(K\cup L)$;
3) дополнительность: $\forall K\subset I:\nu(K)+\nu(I\setminus K)=\nu(I)$.

Пусть $\Upsilon(I)$ — множество всех характеристических функций над I.

Теорема. В пространстве всех вещественных функций над $I:(|I|=n)$ множество $\Upsilon(I)$ является выпуклым конусом размерности не более 2^n-1.

Для нахождения базиса $\Upsilon(I)$ введем важнейшие понятия — простой и простейшей характеристических функций (ХФ).

Характеристическая функция называется простой, если она принимает лишь два значения — 0 и 1. Причем коалиции *K*, для которых $\nu(K)=1$ — выигрывающие, а $\nu(K)=0$ — проигрывающие.

Характеристическая функция называется простейшей, если в простой ХФ выигрывающими являются лишь те коалиции, которые содержат фиксированную непустую коалицию R и обозначаются $\nu_R(K)$. То есть

$$\nu_R(K)=\begin{cases}1, R\subset K\\0, R\not\subset K\end{cases}.$$

Примером простой ХФ является оценка результатов голосования совета безопасности ООН, где выигрывают те коалиции, которые состоят из постоянных пяти членов совета и хотя бы одного непостоянного члена. Все остальные проигрывают.

Простейшие функции возникают, когда в коллективе имеется некое «ядро», голосующее с правом вето, а голоса остальных участников становятся несущественными.

Утверждение. Число различных простейших характеристических функций над I совпадает с числом непустых подмножеств, то есть 2^n-1.

Теорема. Все 2^n-1 простейшие ХФ над I линейно независимы в совокупности.

Следствие. Каждую характеристическую функцию над I можно представить, причем единственным образом, в виде линейной комбинации простейших ХФ.

Теорема. Для любой ХФ имеет место разложение

$$\nu=\sum_{R\subset I}\lambda_R(\nu)\nu_R(K),$$

где $\lambda_R(\nu)=\sum_{K\subset R}(-1)^{|R|-|K|}\nu(K)$.

ХФ ν над I изоморфна ХФ ν' над I', если существует взаимно однозначное отображение $\pi: I\to I'$, такое, что для всякой коалиции K $\nu'(\pi K)=\nu(K)$.

Изоморфизм характеристической функции на себя называют автоморфизмом: $\nu(\pi K)=\nu(K)$.

Игрок i в игре с ХФ v над I — существенный, если найдется такая коалиция K, для которой $v(K)+v(i)<v(K\cup i)$. Игрок j называется несущественным, или болваном, если для всякой коалиции $K: v(K)=v(K\cup i)$.

Коалиция, содержащая всех существенных игроков, называется носителем характеристической функции.

Тройка $\Gamma=\{I,v,A_v\}$ называется классической кооперативной игрой (ККИ). Сформулируем естественные условия для дележа:

1° Индивидуальная рациональность: $x_i \geq v(i)$ для $\forall i \in I$;

2° Коллективная рациональность: $x(I)=v(I)$.

Итак, дележом в условии характеристической функции v над I назовем такой вектор $x: I \to \mathbb{R}^n$, который удовлетворяет условиям 1 и 2.

В теории игр рассматриваются различные принципы оптимальности, которые считаются естественными, однако помимо них конструируются принципы, задаваемые определенными свойствами, которыми они должны обладать. Такой подход принято называть аксиоматическим. Далее рассмотрим один из таких принципов, являющийся интересным как со стороны теории, так и с практической точки зрения. Его принято называть методом справедливого дележа.

Зададимся целью указать способ, ставящий в соответствие каждой игре с заданной характеристической функцией v над множеством игроков $I=\{1,\dots,n\}$ вектор $\Phi(v)=(\Phi_1(v),\dots,\Phi_n(v))$, чьи компоненты описывают «справедливые» выигрыши каждого из игроков. При этом необходимо, чтобы такой вектор оказался дележом в условиях v. Данный вектор принято называть вектором Шепли (ВШ). Сформулируем аксиомы для такого вектора.

Аксиомы Шепли:

1° Аксиома эффективности: пусть I — носитель игры v, тогда

$$\sum_{i\in I}\Phi_i(v)=v(I);$$

2° Аксиома симметрии: для всякого автоморфизма π игры

$$v \quad \Phi_i(v)=\Phi_{\pi i}(v);$$

3° Аксиома агрегации: если v' и v'' — две игры со множеством игроков I, то

$$\Phi_i(v'+v'')=\Phi_i(v')+\Phi_i(v'');$$

4° Аксиома болвана: если игрок i — болван, то

$$\Phi_i(v)=0.$$

Компонента вектора Шепли будет находиться по формуле

$$\Phi_i(v) = \sum_{K|i \in K} \frac{(k-1)!(n-k)!}{n!} \left[v(K) - v(K \setminus i) \right].$$

Если в формуле для вычисления компонент ВШ суммирование проводить по коалициям, не содержащим игрока i, то получим другое представление для ВШ, которое используется так же часто, как и исходное:

$$\Phi_i(v) = \sum_{K|i \notin K} \frac{k!(n-k-1)!}{n!} \left[v(K \cup i) - v(K) \right].$$

Применение вектора Шепли в регрессионном анализе

Опишем процедуру применения ВШ в регрессионном анализе. Для начала необходимо задать характеристическую функцию. Для этого поможет значение коэффициента детерминации R^2. Будем использовать его как меру качества модели.

Зададим характеристическую функцию. Пусть значение характеристической функции в модели без предикторов будет равно 0: $v(0) = 0$. Значение ХФ для модели с лишь j-м предиктором, будет равняться $R_j^{\,2}$, то есть коэффициенту детерминации для модели $y_i = b_j x_{ij}$. Аналогично в случае нескольких предикторов, например значение $v(x_j x_k) = R_{jk}^{\,2}$ для модели $y_i = b_j x_{ij} + b_k x_{ik}$. И т. д. Окончательно характеристическая функция для модели со всеми предикторами будет равняться $v(x_1 x_2 ... x_n) = R_{all}^{\,2} = R^2$, то есть коэффициенту детерминации для исходной модели.

Заметим, что, помимо классического способа нахождения коэффициента детерминации, можно воспользоваться и альтернативным. Для нормированной модели выполняется следующее: $R^2 = 1 - S^2$.

Тогда воспользуемся данным фактом, зная, что

$$S^2 = Y^t Y - 2Y^t Xb + b^t X^t Xb = Y^t Y - b^t r + b^t Cb.$$

Поскольку отклик пронормирован, то $Y^t Y = 1$. Тогда:

$$R^2 = 1 - S^2 = b^t r + b^t Cb = \sum_j b_j (2r - Cb)_j.$$

Вспоминая систему нормальных уравнений $Cb = r$, получим

$$R^2 = \sum_j b_j (2r - Cb)_j = \sum_j b_j r_j.$$

Величину под знаком суммы будем называть $NEF_j = b_j r_j$ (net effects).

Соответственно, коэффициент детерминации будет равен сумме всех NEF.

Далее находятся коэффициенты, или веса, для ВШ по формуле

$$\gamma_i(K) = \frac{(k-1)!(n-k)!}{n!} = \frac{1}{k \cdot C_n^k},$$

где k — количество предикторов в рассматриваемой модели, а n — общее количество предикторов в исходной модели.

Затем находятся значения ВШ по формуле

$$SV_i = \sum_{K|i \in K} \frac{(k-1)!(n-k)!}{n!} \left[v(K) - v(K \setminus i) \right].$$

Например, в модели с n предикторами значение ВШ для предиктора x_1 будет иметь следующий вид:

$$SV_1 = \frac{1}{n}\left(R_1^{\ 2}\right) + \frac{1}{n(n-1)}\left(R_{12}^{\ 2} - R_2^{\ 2} + R_{13}^{\ 2} - R_3^{\ 2} + \ldots + R_{1n}^{\ 2} - R_n^{\ 2}\right) + \ldots +$$

$$+\frac{1}{n(n-1)}\left(R_{13\ldots n}^{\ 2} - R_{34\ldots n}^{\ 2} + R_{124\ldots n}^{\ 2} - R_{24\ldots n}^{\ 2} + \ldots + R_{12\ldots n-1}^{\ 2} - R_{23\ldots n-1}^{\ 2}\right) +$$

$$+\frac{1}{n}\left(R_{all}^{\ 2} - R_{234\ldots n}^{\ 2}\right).$$

Громоздкость формулы сподвигает на переобозначение — введем коэффициенты полезности [11].

Пусть $R_i^{\ 2} = U_i, R_{ij}^{\ 2} - R_j^{\ 2} = U_{ij}$ и т.д. $R_{all}^{\ 2} - R_{234\ldots n}^{\ 2} = U_{all}$.

Соответственно, предыдущую формулу можно переписать:

$$SV_1 = \frac{1}{n}U_1 + \frac{1}{n(n-1)}\sum_j U_{1j} + \ldots + \frac{1}{n(n-1)}\left(U_{134\ldots n} + U_{124\ldots n} + \ldots + U_{123\ldots n-1}\right) + \frac{1}{n}U_{all}.$$

Легко проверить, что $\sum_{j=1}^{n} SV_j = R^2$.

Можно сделать вывод о возможности рассмотрения компонент вектора Шепли как значения NEF [9]. Будем называть их SV net effects. На самом деле они являются оценками NEF, полученными усреднением по всем возможным моделям с различным числом предикторов.

В отличие от стандартных NEF, аналог Шепли обладает следующими свойствами:

1) все SV_j положительны, что следует из построения вектора Шепли и, соответственно, более репрезентативны и легко представимы в графическом виде;

2) так как они являются значениями NEF, усредненными по всевозможным линейным моделям с различным количеством предикторов, то воздействие мультиколлинеарности на них сведено к минимуму.

Подведем итоги. Значение Шепли для каждого предиктора является мерой его полезности или предельным ожидаемым вкладом, усредненным по всевозможным моделям, включающим данный предиктор.

Соответственно, сумма таких вкладов образует коэффициент детерминации модели. Так как компоненты вектора Шепли являются в какой-то степени долями R^2, то они означают значимость каждого предиктора в модели.

Окончательно формулу можно переписать в виде

$$SV_{x_1} = \frac{1}{n}\left(R_{all}{}^2\right) + \frac{1}{n-1}\left(R_{x_1}{}^2 - \overline{R_1{}^2}\right) + \frac{1}{n-2}\left(\overline{R_{x_1*}{}^2} - \overline{R_2{}^2}\right) + \ldots +$$
$$+ \frac{1}{\left(n-(n-1)\right)}\left(\overline{R_{x_1*\ldots*}{}^2} - \overline{R_{n-1}{}^2}\right).$$

где $\overline{R_{x_1*}{}^2}$ — среднее значение для всех моделей, состоящих из 2 предикторов, один из которых x_1; $\overline{R_j^2}$ — среднее значение для всех моделей, состоящих из j-предикторов.

Первое слагаемое показывает вклад, который предиктор x_1 вносит в модель со всеми предикторами. Второе слагаемое показывает разность между значением R^2 для модели лишь с одним предиктором x_1 и средним значением R^2 для всех моделей с одним предиктором. И т. д.

Исходя из того, что $\sum_{j=1}^{n} SV_j = R^2$, построим оценки, аналогичные МНК.

Оценки вида $a_j = \frac{SV_j}{r_j}$, где SV_j — компонента ВШ для j-го предиктора, а r_j — j-й элемент вектора X^tY, или, по сути, коэффициент корреляции между x_j и y, будем называть оценками методом Шепли, или оценками Шепли.

Перейдем от $SV_j = a_j r_j$ к $SV_j = a_j\left(2r - Ca\right)_j$. Для заранее найденных значений Шепли данные выражения представляют собой систему из n квадратичных уравнений, которая может быть решена относительно a_j методами нелинейной оптимизации выражения:

$$F = \sum_j \left[SV_j - a_j\left(2r - Ca\right)_j\right]^2 = \sum_j \left[SV_j - 2a_j r_j + a_j \sum_k r_{jk} a_k\right]^2.$$

Коэффициенты $a_j = \frac{SV_j}{r_j}$ могут быть использованы как начальное приближение a_j^0 в минимизационной процедуре [6; 7].

Полученные в результате данной процедуры коэффициенты — коэффициенты скорректированной регрессии. Такие коэффициенты не подвержены вредным действиям мультиколлинеарности и имеют интерпретабельные знаки и значения.

Заключение

В статье было рассмотрено применение инструмента из теории кооперативных игр, а именно вектора Шепли, для оценки коэффициентов регрессии и относительной значимости предикторов в модели линейной регрессии.

Список литературы

1. *Воробьев Н.Н.* Теория игр для экономистов-кибернетиков. М., 1985.
2. *Вучков И., Бояджиева Л., Солаков Е.* Прикладной линейный регрессионный анализ. М., 1987.
3. *Демиденко Е.З.* Линейная и нелинейная регрессия. М., 1981.
4. *Домбровский В.В.* Эконометрика. М., 2004.
5. *Петросян Л.А., Зенкевич Н.А., Шевкопляс Е.В.* Теория игр. СПб., 2012.
6. *Эфрон Б.* Нетрадиционные методы многомерного статистического анализа. М., 1988.
7. *Breiman L.* Bagging predictors // Machine learning. 1994. Vol. 26, P. 123—140.
8. *Conklin M., Powaga K., Lipovetsky S.* Customer satisfaction analysis: identification of key drivers // European Journal of Operational Research 2004. Vol. 154. P. 819—827.
9. *Draper N., Smith H.* Applied regression analysis. N. Y., 1998.
10. *James M., Stein C.* Estimation with quadratic loss // Proceedings of the Fourth Berkley Symposium. UCP, 1961.
11. *Harris R.A* Primer of Multivariate Statistics. N.Y. ; L., 1975.
12. *Lipovetsky S., Conklin M.* Analysis of regression in game theory approach // 2001. Appl. Stochastic Models Bus. Ind. Vol. 17. P. 319—330.
13. *Mason Ch. H., Perreault Jr. W.D.* Collinearity, power, and interpretation of multiple regression analysis // Journal of Marketing Research. 1991. Vol. 28. P. 268—280.
14. *Montgomery D., Peck E., Vining G.* Introduction to Linear Regression Analysis. John Wiley & Sons, 2012.
15. *The Shapley* Value: Essays in Honor of Lloyd S. Shapley / ed. by A.E. Roth. Cambridge, 1988.

Об авторах

Артём Михайлович Михеенко — магистрант, Балтийский федеральный университет им. И. Канта, Россия.
E-mail: Artem-miheenko9@yandex.ru

Дарья Сергеевна Савич — магистрант, Балтийский федеральный университет им. И. Канта, Россия.
E-mail: Darja.savich@yandex.ru

The authors

Artyom M. Mikheyenko, Master's Student, Immanuel Kant Baltic Federal University, Russia.
E-mail: Artem-miheenko9@yandex.ru

Darja S. Savich, Master's Student, Immanuel Kant Baltic Federal University, Russia.
E-mail: Darja.savich@yandex.ru