

УДК 519.2, 51-7, 311, 303.732

Е. Ю. Чудина, Т. В. Жмыхова

Государственное образовательное учреждение высшего профессионального образования
«Донбасская национальная академия строительства и архитектуры» (ГОУ ВПО ДОННАСА)
г. Макеевка-23, ул. Державина, 2

О ПРИМЕНЕНИИ МЕТОДА ИДЕАЛЬНОЙ ТОЧКИ В УСЛОВИЯХ МНОГОФАКТОРНОЙ ЛИНЕЙНОЙ РЕГРЕССИИ ПРИ ОТСУТСТВИИ МУЛЬТИКОЛЛИНЕАРНОСТИ ФАКТОРНЫХ ПРИЗНАКОВ

K. Y. Chudina, T. V. Zhmykhova

State educational institution of higher professional education
«Donbas National Academy of Civil Engineering and Architecture» (DONNACEA)
Makeevka-23, st. Derzhavina, 2

ON THE APPLICATION OF THE IDEAL POINT METHOD IN THE CONDITIONS OF MULTIVARIATE LINEAR REGRESSION IN THE ABSENCE OF MULTICOLLINEARITY OF FACTOR FEATURES

В работе рассматривается математическая модель взаимосвязи уровня учебной успеваемости студентов первого года обучения в учреждении высшего профессионального образования с уровнем их успеваемости в средней школе (общей и по профильному предмету). Предложена модель оптимизации успеваемости студентов первого года обучения технического вуза по профильному предмету, основанная на статистических данных.

Ключевые слова: уровень успеваемости, задача оптимизации, условный экстремум двух линейных функций, метод идеальной точки, множество Парето, регрессионный анализ.

The mathematical model of the relationship between the level of achievement of students of the first year education in an institution of higher professional education with the level of their performance in secondary school (general and in a specialized subject) is considered. A model for optimizing the academic performance of first-year students of a technical university in a specialized subject based on statistical data is proposed.

Key words: academic performance level, optimization problem, conditional extremum of two linear functions, ideal point method, Pareto set, regression analysis.

Введение

Изучению проблемы повышения успеваемости обучающихся в учреждениях профессионального образования посвящен ряд работ, в которых исследовались направления, методы и формы работы по повышению успеваемости [1], факторы, влияющие на снижение успеваемости [2]; выделялись основные группы причин неуспеваемости [3] и пр. Однако мало научных изысканий ведется в области математического моделирования успеваемости обучающихся. Отдельные исследования используют, как правило, только статистический аппарат. Исследования, проводимые в данном направлении, имеют цель изучения факторов, влияющих на успеваемость студентов, среди которых можно выделить личностные качества студента, процессы адаптации, социально-психологический климат коллектива, посещаемость занятий и интерес к занятиям по изучаемому предмету, педагогический талант преподавателя, трудовая занятость в период учебы, время, затрачиваемое на самостоятельное обучение. Этим вопросам посвящен ряд работ [4-7]. Недостатком данных работ является то, что все они направлены на изучение внутренних факторов с целью перенаправления совместных усилий факультетов, кафедр и преподавателей на достижение студентами высокого уровня усвоения знаний, а изучение влияния внешних факторов минимально или отсутствует вообще.

Авторами была поставлена задача построения математической модели взаимосвязи успеваемости студентов инженерных специальностей первого года обучения по профильному предмету в учреждении высшего профессионального образования, общей успеваемости абитуриентов и успеваемости по профильному предмету при получении общего среднего образования; выявления уровня успеваемости абитуриентов, оптимального для максимизации успеваемости студентов первого года обучения по профильному предмету.

Постановка задачи. Согласно [8], [9], конкурсный балл для абитуриентов, поступающих на обучение по образовательным программам бакалавриата (специалитета), рассчитывается как сумма среднего балла (по 100-балльной шкале) документа о полученном среднем общем образовании, оценки из сертификата государственной итоговой аттестации (ГИА) (по 100-балльной шкале) по русскому языку (обязательный предмет) и оценки по одному профильному конкурсному предмету (предмет на выбор) или балла дополнительного вступительного испытания, умноженных на весовые коэффициенты, и дополнительных баллов. Устанавливаются следующие весовые коэффициенты: средний балл аттестата: 0,2; оценка из сертификата ГИА по русскому языку: 0,3; оценка из сертификата ГИА по профильному предмету или балл дополнительного вступительного испытания: 0,5.

Таким образом, конкурсный балл абитуриента рассчитывается как:

$$X = 0,2X_1 + 0,5X_2 + 0,3X_3 + X_4, \quad (1)$$

где X – конкурсный балл абитуриента при поступлении в вуз, X_1 – уровень общей успеваемости абитуриентов (средний балл аттестата по 100-балльной шкале), X_2 – уровень успеваемости абитуриентов по профильному предмету (балл ГИА по профильному предмету), X_3 – уровень успеваемости абитуриентов по русскому языку (балл ГИА по русскому языку), X_4 – дополнительные баллы, начисляемые

абитуриентам, награжденным золотой или серебряной медалью (7 или 5 баллов соответственно); имеющим диплом о среднем профессиональном образовании с отличием (5 баллов); абитуриентам, окончившим Республиканский архитектурно-строительный лицей-интернат ГОУ ВПО «ДонНАСА» (5 баллов).

Была поставлена задача максимизации конкурсного балла абитуриента (как условия поступления в учреждение высшего профессионального образования) и максимизации уровня успеваемости обучающихся по профильному предмету при заданных ограничениях (с учетом 200-балльной шкалы):

$$0 \leq X_1 \leq 200, \quad 160 \leq X_2 \leq 200, \quad 160 \leq X_3 \leq 200, \quad 0 \leq X_4 \leq 7. \quad (2)$$

Основная часть

Пусть Y – уровень успеваемости студентов вуза по профильному предмету (математике). Составим уравнение множественной линейной регрессии:

$$Y = a_1 X_1 + a_2 X_2 + a_3 X_3 + a_4 X_4 + a_5, \quad (3)$$

где коэффициенты a_1, a_2, a_3, a_4, a_5 находятся эмпирически.

В статье рассмотрены данные об успеваемости 50 обучающихся ГОУ ВПО «Донбасская национальная академия строительства и архитектуры» за 2018 – 2019 уч.г. инженерных направлений подготовки, а именно «Водоснабжение и водоотведение», «Инженерная защита окружающей среды», «Информационно-стоимостной инжиниринг» и «Менеджмент строительных организаций» (табл. 1), отметим, что в задачи исследования не входило изучение влияния форс-мажорных факторов, таких как пандемия COVID-19.

Таблица 1 – Данные о баллах абитуриентов по 200-балльной шкале

№	Русский язык	Математика	Аттестат	Доп. балл	Математика (итог)
1	195	173	198	0	90
2	162	124	160	0	68
3	170	147	176	0	75
...
49	199	200	171,9	0	75
50	129	187	168,5	0	75

Эмпирическое уравнение регрессии имеет вид:

$$Y = 31,3395 - 0,0083 X_1 - 0,0646 X_2 + 0,3323 X_3 + 0,0813 X_4. \quad (4)$$

Парные коэффициенты корреляции между факторными и результативным признаками составили:

$$r_{YX_1} = -0,0980; \quad r_{YX_2} = 0,0348; \quad r_{YX_3} = 0,5820; \quad r_{YX_4} = -0,1040. \quad (5)$$

Таким образом, наиболее сильная связь результативного признака Y – уровня успеваемости студентов вуза по математике – наблюдается с факторным признаком X_3 – средним баллом аттестата.

$$\begin{aligned} r_{X_1 X_2} &= 0,4660; & r_{X_1 X_3} &= 0,0428; & r_{X_1 X_4} &= 0,0802; \\ r_{X_2 X_3} &= 0,3552; & r_{X_2 X_4} &= 0,2123; & r_{X_3 X_4} &= -0,1205. \end{aligned} \quad (6)$$

Наиболее сильная связь между факторными признаками наблюдается между оценками по математике и русскому языку в средней школе, а также между оценкой по математике и средним баллом аттестата. Связь умеренная и прямая, то есть увеличение значения одного из этих факторов ведет к увеличению другого. Примечательно, что при этом между значениями среднего балла аттестата и дополнительными баллами абитуриента присутствует обратная связь.

Все парные коэффициенты корреляции меньше 0,7, что говорит об отсутствии мультиколлинеарности факторных признаков. Критерий Стьюдента показал, что коэффициенты корреляции r_{YX_2} , r_{YX_3} статистически значимы. Связь между признаками Y и X_3 наиболее выражена, поэтому выбираем X_3 в качестве первого фактора. Частные коэффициенты корреляции при введении в модель фактора X_3 составили:

$$r_{YX_3/X_1} = 0,589; \quad r_{YX_3/X_2} = 0,609; \quad r_{YX_3/X_4} = 0,576. \quad (7)$$

Таким образом, вторым значимым фактором целесообразно выбрать X_2 . Эти выводы подтверждают высказанное нами ранее предположение о том, что успеваемость обучающихся в техническом вузе коррелирует именно с оценкой по профильному предмету (математике) и средним баллом аттестата [10]. В этом случае эмпирическое уравнение регрессии примет вид

$$Y = 31,062 - 0,0711X_2 + 0,3336X_3. \quad (8)$$

Средняя ошибка аппроксимации составила 0,078, что говорит о соответствии модели. Критерий Фишера также показывает, что уравнение регрессии статистически надежно:

$$F = 13,927 > F_{\text{крит}}(2;47) = 3,15. \quad (9)$$

Итак, исключая незначимые факторы X_1 и X_4 , сформулируем следующую задачу оптимизации [11], [12].

$$\begin{cases} L_1 = 0,2X_2 + 0,5X_3 \rightarrow \max; \\ L_2 = -0,0711X_2 + 0,3336X_3 \rightarrow \max; \\ 100 \leq X_2 \leq 200, \quad 160 \leq X_3 \leq 200. \end{cases} \quad (10)$$

Областью допустимых решений в двумерной системе координат OX_2X_3 будет прямоугольник с вершинами $A(100;160)$, $B(200;160)$, $C(200;200)$, $D(100;200)$. Координаты каждой точки плоскости OX_2X_3 подвергнем линейным преобразованиям $L_1 = 0,2X_2 + 0,5X_3$ и $L_2 = -0,0711X_2 + 0,3336X_3$. Получим систему координат OL_1L_2 , при этом, в силу линейности преобразований, прямоугольная система координат OX_2X_3 перейдет в прямоугольную систему координат OL_1L_2 , а прямоугольник $ABCD$ перейдет в многоугольник $A_1B_1C_1D_1$, вершины которого имеют соответственно координаты:

$$A_1(100;42,266), \quad B_1(120;39,156), \quad C_1(140;52,5), \quad D_1(120;59,61) \text{ (рис. 1)}.$$

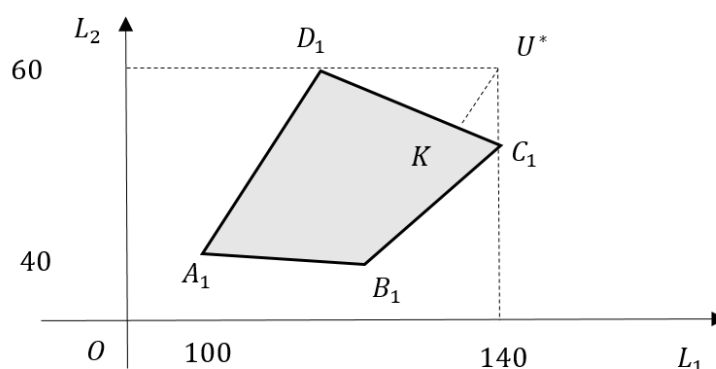


Рисунок 1 – Область допустимых решений в пространстве критериев и множество Парето

Множеством Парето является множество точек отрезка C_1D_1 . В качестве точки утопии выбираем точку $U(140;59,61)$, отвечающую наибольшему значению L_1 , L_2 . Уравнение прямой C_1D_1 : $L_2 = 102,27 - 0,3555L_1$, уравнение перпендикуляра к C_1D_1 из точки утопии: $L_2 = -334,202 + 2,813L_1$. Тогда оптимальная точка: $K(137,754; 53,299)$.

Оптимальные значения X_2 , X_3 найдем из системы уравнений:

$$\begin{cases} 0,2X_2 + 0,5X_3 = 137,754; \\ -0,0711X_2 + 0,3336X_3 = 53,299. \end{cases} \quad (11)$$

Получим оптимальные решения: $X_2 = 188,769$, $X_3 = 200$.

Выводы

Результаты построения математической модели успеваемости обучающихся при изучении профильного предмета (математики) в техническом вузе показали, что ее уровень зависит в большей мере от среднего балла аттестата абитуриента, чем от уровня подготовки по этой дисциплине. Такие факторы, как оценка по обязательной непрофильной дисциплине (русскому языку) и дополнительные баллы абитуриента, являются статистически малозначимыми. Были получены оптимальные значения оценки абитуриента при изучении математики в средней школе и среднего балла аттестата, при которых достигается максимум двух целевых функций – уровня подготовки абитуриента, выражающийся в максимально возможной сумме баллов при поступлении в вуз, и в дальнейшем – наиболее высоком уровне усвоения математики как профильной дисциплины в техническом вузе.

Список литературы

1. Шибает В. П. Система работы по повышению успеваемости студентов [Электронный ресурс] / В. П. Шибает, Л. М. Шибаета // Мир науки, культуры, образования. – № 4 (41). – 2013. – Режим доступа: <http://cyberleninka.ru/article/n/sistemaraboty-po-povysheniyu-uspevaemosti-studentov> (дата обращения: 29.03.21).
2. Пономарчук Е. С. Заинтересованность студентов в учебе. [Электронный ресурс] // Е. С. Пономарчук, Н. В. Шамаева // Материалы III студенческой конференции БГТУ им. В. Г. Шухова «Образование, наука, производство». – Режим доступа: www.cs-alternativa.ru/text/2040/2 (дата обращения: 29.03.18).

3. Бабанский Ю. К. Оптимизация процесса обучения (Общедидактический аспект) [Текст] / Ю. К. Бабанский. – М.: Педагогика, 1977. – 253 с.
4. Кошелева Г. В. Факторы, влияющие на успеваемость студентов [Текст] // Г. В. Кошелева, Ю. Ю. Фионова // Актуальные направления научных исследований XXI века: теория и практика. 2015. – № 7–4 (18–4). – С. 331–333.
5. Соловьев В. Н. Влияние различных двигательных режимов на успеваемость и адаптацию студентов к учебному процессу в вузе [Текст] / В. Н. Соловьев // Успехи современного естествознания. – 2004. – № 3. – С. 92–93.
6. Гранков М. В. Анализ и кластеризация основных факторов, влияющих на успеваемость учебных групп вуза [Электронный ресурс] / М. В. Гранков, В. М. Аль-Габри, М. Ю. Горлова // Инженерный вестник Дона. – 2016. – № 4. – Режим доступа : <http://ivdon.ru/magazine/archive/n4y2016/3775> (Дата обращения: 22.05.2021).
7. Городецкая И. В. Оценка факторов, положительно и отрицательно влияющих на успеваемость студентов» [Электронный ресурс] // И. В. Городецкая, В. Г. Захаревич. – Режим доступа: <https://cyberleninka.ru/article/n/otsenka-faktorov-polozhitelno-i-otritsatelno-vliyayuschih-na-uspevaemost-studentov> (дата обращения: 16.14.2021)/.
8. Приказ Министерства образования и науки Донецкой Народной Республики от 17.11.2017 г. №1209.
9. Порядок приема на обучение в образовательные организации высшего профессионального образования Донецкой Народной Республики на 2018/2019 учебный год [Электронный ресурс]. – Режим доступа: http://new.donnasa.ru/docs/priemnaya_komissiya/pravila_priema/pril._prikaz_mon_dnr_%E2%84%96_1209_ot_17.11.2017g..pdf (дата обращения: 21.03.18).
10. Чудина Е. Ю. Исследование успеваемости студентов первого года обучения методами оптимизации [Электронный ресурс] / Е. Ю. Чудина, Т. В. Жмыхова // Международная конференция преподавателей, сотрудников, докторантов и аспирантов. Соврем. проблемы теории вероятностей и мат. статистики. Актуарная и финансовая математика, посвященная 100-летию со дня рождения И.И. Гихмана. 24 мая – 26 мая 2018 года, г. Донецк. – 56 с. – С. 22–23.
11. Дубов Ю. А. Многокритериальные модели формирования и выбора вариантов систем [Текст] / Ю. А. Дубов, С. И. Травкин, В. Н. Якимец. – М.: Наука, 1986. – 296 с.
12. Моисеев Н. Н. Математические задачи системного анализа [Текст] / Н. Н. Моисеев – М. : Наука, 1981. – 488 с.

References

1. Shibaev V. P., Shibaeva L. M. Sistema raboty po povysheniyu uspevayemosti studentov [The system of work on improving the academic performance of students]. *Mir nauki, kul'tury, obrazovaniya* [The world of science, culture, education], No. 4 (41), 2013, Access mode: <http://cyberleninka.ru/article/n/sistemaraboty-po-povysheniyu-uspevaemosti-studentov> (date of application: 29.03.21).
2. Ponomarchuk E. S., Shamaeva N. V. Zainteresovannost' studentov v uchebe [Students ' interest in studying] *Materialy III studencheskoy konferentsii BGTU im. V. G. Shukhova «Obrazovaniye, nauka, proizvodstvo»* [Materials of the III student conference of V. G. Shukhov BSTU "Education, science, production"]. Access mode: www.cs-alternativa.ru/text/2040/2 (accessed: 29.03.18).
3. Babansky Yu. K. *Optimizatsiya protsessa obucheniya (Obshchedidakticheskiy aspekt)* [Optimization of the learning process (General didactic aspect)], М., "Pedagogy", 1977, 253 p.
4. Kosheleva G. V. Fionova Yu. Yu. Faktory, vliyayushchiye na uspevayemost' studentov [Factors affecting the academic performance of students] *Aktual'nyye napravleniya nauchnykh issledovaniy XXI veka: teoriya i praktika* [Current directions of scientific research of the XXI century: theory and practice], 2015, No. 7-4 (18-4), pp. 331-333.
5. Solovyov V. N. Vliyaniye razlichnykh dvigatel'nykh rezhimov na uspevayemost' i adaptatsiyu studentov k uchebnomu protsessu v vuze [The influence of various motor modes on the academic performance and adaptation of students to the educational process at the university] *Uspekhi sovremennogo yestestvoznaniya* [Uspekhi sovremennogo naturalnozovaniya, 2004, No. 3, pp. 92-93.
6. Grankov M. V., Al-Gabri V. M., Gorlova M. Yu. naliz i klasterizatsiya osnovnykh faktorov, vliyayushchikh na uspevayemost' uchebnykh grupp vuza [Analysis and clustering of the main factors affecting the academic performance of university study groups]. *Inzhenernyy vestnik Dona* [Engineering Bulletin of the Don], 2016, No. 4. Access mode: <http://ivdon.ru/magazine/archive/n4y2016/3775> (Accessed: 22.05.2021).

7. Gorodetskaya I. V., Zakharevich V. G. *Otsenka faktorov, polozhitel'no i otritsatel'no vliyayushchikh na uspevayemost' studentov* [Assessment of factors that positively and negatively affect the academic performance of students"] Access mode: <https://cyberleninka.ru/article/n/otsenka-faktorovpolozhitelno-i-otritsatelno-vliyayuschih-na-uspevaemost-studentov> (accessed: 16.14.2021)/.
8. *Prikaz Ministerstva obrazovaniya i nauki Donetskoy Narodnoy Respubliki ot 17.11.2017 g. №1209*. [Order of the Ministry of Education and Science of the Donetsk People's Republic No. 1209 of 17.11.2017].
9. *Poryadok priyema na obucheniye v obrazovatel'nyye organizatsii vysshego professional'nogo obrazovaniya Donetskoy Narodnoy Respubliki na 2018/2019 uchebnyy god* [The procedure for admission to training in educational organizations of higher professional education of the Donetsk People's Republic for the 2018/2019 academic year]. \ Access mode: http://new.donnasa.ru/docs/priemnaya_komissiya/pravila_priema/pril._prikaz_mon_dnr_%E2%84%96_1209_ot_17.11.2017g..pdf (accessed: 21.03.18).
10. Chudina E. Yu., Zhmykhova T. V. Issledovaniye uspevayemosti studentov pervogo goda obucheniya metodami optimizatsii [Research of the academic performance of students of the first year of study by optimization methods]. *Mezhdunarodnaya konferentsiya pre-podavateley, sotrudnikov, doktorantov i aspirantov. Sovrem. problemy teorii veroyatnostey i mat. statistiki. Aktuarnaya i finansovaya matematika, posvyashchennaya 100-letiyu so dnya rozhdeniya I.I. Gikhmana* [International Conference of teachers, Staff, doctoral students and postgraduates. We'll lie. problems of probability theory and math. statistics. Actuarial and financial mathematics, dedicated to the 100th anniversary of the birth of I. I. Gihman]. May 24-May 26, 2018, Donetsk. 56 p., p. 22-23.
11. Dubov Yu. A., Travkin S.I., Yakimets V. N. *Mnogokriterial'nyye modeli formirovaniya i vybora variantov sistem* [Multicriteria models of formation and choice of system variants], M., Nauka, 1986, 296 p.
12. Moiseev N. N. *Matematicheskiye zadachi sistemnogo analiza* [Mathematical problems of system analysis], M., Nauka, 1981, 488 p.

RESUME

K.Y. Chudina, T.V. Zhmykhova

On the application of the ideal point method in the conditions of multivariate linear regression in the absence of multicollinearity of factor features

The article considers a mathematical model of the relationship between the level of academic performance of students of the first year of study in an institution of higher professional education with the level of their performance in secondary school (general and in a specialized subject).

A model of optimizing the academic performance of students of a technical university of the first year of study in a specialized subject, based on statistical data, is proposed.

Based on statistical data, the empirical dependence of the effective factors is determined, significant factors are identified. The objective functions of the students' academic performance level are introduced and their maximum is found under the given constraints by the optimal point method.

The results of constructing a mathematical model of students' academic performance in studying a specialized subject (mathematics) at technical university showed that its level depends more on the average score of the applicant's certificate than on the level of training in this discipline. Such factors as the assessment of a mandatory non-core discipline (Russian language) and the applicant's additional points are statistically insignificant. The optimal values of the applicant's assessment when studying mathematics in secondary school and the average score of the certificate were obtained, at which a maximum of two target functions is achieved – the level of preparation of the applicant and the level of academic performance when studying mathematics as a profile discipline in a technical university.

РЕЗЮМЕ

Е. Ю. Чудина, Т. В. Жмыхова

О применении метода идеальной точки

в условиях многофакторной линейной регрессии

при отсутствии мультиколлинеарности факторных признаков

В статье рассмотрена математическая модель взаимосвязи уровня учебной успеваемости студентов первого года обучения в учреждении высшего профессионального образования с уровнем их успеваемости в средней школе (общей и по профильному предмету).

Предложена модель оптимизации успеваемости обучающихся технического вуза первого года обучения по профильному предмету, основанная на статистических данных.

На основании статистических данных определена эмпирическая зависимость результативных факторов, выявлены значимые факторы. Введены целевые функции уровня успеваемости обучающихся и найден их максимум при заданных ограничениях методом оптимальной точки.

Результаты построения математической модели успеваемости обучающихся при изучении профильного предмета (математики) в техническом вузе показали, что уровень успеваемости зависит в большей мере от среднего балла аттестата абитуриента, чем от уровня подготовки по этой дисциплине. Такие факторы, как оценка по обязательной непрофильной дисциплине (русскому языку) и дополнительные баллы абитуриента, являются статистически малозначимыми. Были получены оптимальные значения оценки абитуриента при изучении математики в средней школе и среднего балла аттестата, при которых достигается максимум двух целевых функций – уровня подготовки абитуриента и уровня успеваемости при изучении математики как профильной дисциплины в техническом вузе.

Статья поступила в редакцию 14.06.2021.