

DOI: 10.51790/2712-9942-2022-3-4-04

МНОЖЕСТВЕННОЕ L_v -ОЦЕНИВАНИЕ ЛИНЕЙНЫХ РЕГРЕССИОННЫХ МОДЕЛЕЙ**С. И. Носков^a, М. П. Базилевский^b***Иркутский государственный университет путей сообщения, г. Иркутск, Российская Федерация*^a ORCID: <http://orcid.org/0000-0003-4097-2720>, sergey.noskov.57@mail.ru^b ORCID: <http://orcid.org/0000-0002-3253-5697>, mik2178@yandex.ru

Аннотация: для оценки моделей множественной линейной регрессии существует много различных математических методов: наименьших квадратов, модулей, антиробастного оценивания, L_v -оценивания, множественного оценивания. Целью данной работы является обобщение указанных методов оценивания единой функцией потерь. Сначала была сформулирована задача оценивания, в которой в качестве критериев минимизации выступают критерии для антиробастного и L_v -оценивания. Недостатком сформулированной задачи является то, что для ее численного решения затруднительно определять начальные значения параметров, поскольку переменные могут иметь разные масштабы. Кроме того, функция потерь для этой задачи является неоднородной, что также затрудняет процесс оценивания. Для решения этих проблем введен новый критерий, равный критерию антиробастного оценивания, возведенному в степень v . С помощью него и функции потерь для L_v -оценивания сформулирована задача множественного L_v -оценивания. Функционал этой задачи является однородным, поэтому для проведения множественного L_v -оценивания целесообразно нормировать исходные переменные и переходить к оценкам стандартизированной линейной регрессии. Предложен алгоритм, по которому рекомендуется проводить множественное L_v -оценивание. В результате проведения множественного L_v -оценивания формируется множество, содержащее оценки линейной регрессии, полученные как известными методами, так и новыми. Правильный выбор наилучших из полученного множества оценок пока остается открытой научной задачей. С помощью предложенного множественного L_v -оценивания успешно решена задача моделирования железнодорожных пассажирских перевозок Иркутской области.

Ключевые слова: регрессионная модель, метод наименьших модулей, метод наименьших квадратов, метод антиробастного оценивания, метод L_v -оценивания, метод множественного оценивания, метод множественного L_v -оценивания.

Для цитирования: Носков С. И., Базилевский М. П. Множественное L_v -оценивание линейных регрессионных моделей. *Успехи кибернетики*. 2022;3(4):32–40. DOI: 10.51790/2712-9942-2022-3-4-04.

MULTIPLE L_v -ESTIMATION OF LINEAR REGRESSION MODELS**S. I. Noskov^a, M. P. Bazilevskiy^b***Irkutsk State Transport University, Irkutsk, Russian Federation*^a ORCID: <http://orcid.org/0000-0003-4097-2720>, sergey.noskov.57@mail.ru^b ORCID: <http://orcid.org/0000-0002-3253-5697>, mik2178@yandex.ru

Abstract: there are many methods for estimating multiple linear regression models: ordinary least squares, least absolute deviations, anti-robust estimation, L_v -estimation, and multiple estimations. The purpose of this work is to generalize these methods by a loss function. First, an estimation problem was formulated where the minimization criteria are the anti-robust and L_v -estimations. The disadvantage of this problem statement is that it is difficult to determine the initial values of the parameters for a numerical solution, since the variables may have different scales. Besides, the loss function is non-uniform, which also complicates the estimation. To solve these problems, we introduced a new criterion, equal to the anti-robust estimation criterion raised to the power v . We stated the problem of multiple L_v -estimation using the new criterion and the loss function. The functional of this problem is homogeneous, therefore, for multiple L_v -estimations, it is advisable to normalize the initial variables and then apply the standardized linear regression estimates. We also developed an algorithm for multiple L_v -estimations. A result of such estimations is a set containing linear regression estimates obtained both by the existing and new methods. The optimal choice of the best estimates from the set of estimates remains an open problem. We successfully simulated the passenger railway traffic in the Irkutsk region with the proposed multiple L_v -estimations.

Keywords: regression model, least absolute deviations, ordinary least squares, anti-robust estimation, L_v -estimation, multiple estimations, multiple L_v -estimation method.

Cite this article: Noskov S. I., Bazilevskiy M. P. Multiple L_v -estimation of Linear Regression Models. *Russian Journal of Cybernetics*. 2022;3(4):32–40. DOI: 10.51790/2712-9942-2022-3-4-04.

Введение

В настоящее время одним из эффективных инструментов математического моделирования является регрессионный анализ [1, 2]. Простейшей регрессионной моделью считается линейная регрессия вида:

$$y_i = \alpha_0 + \sum_{j=1}^m \alpha_j x_{ij} + \varepsilon_i, \quad i = \overline{1, n}, \quad (1)$$

где n — объем выборки; m — число объясняющих переменных; y_i — i -е значение объясняемой переменной; x_{ij} — i -е значение j -й объясняющей переменной; $\alpha_0, \alpha_1, \dots, \alpha_m$ — неизвестные параметры; ε_i — i -я ошибка аппроксимации.

Существуют различные методы нахождения неизвестных параметров линейной регрессии (1): метод максимального правдоподобия, моментов, наименьших квадратов и модулей, антиробастного и множественного оценивания и т. д. При этом вопрос о том, какой из существующего арсенала метод лучше подходит для оценки модели (1), остается нерешенным. Большинство исследователей в своих научных работах отдает предпочтение методу наименьших квадратов (МНК) [3, 4], состоящему в решении задачи:

$$J_2(\alpha) = \sum_{i=1}^n \varepsilon_i^2 \rightarrow \min. \quad (2)$$

Распространенность МНК объясняется простотой решения задачи (2), сводящейся к решению системы линейных алгебраических уравнений.

Однако эффективность МНК резко снижается тогда, когда обрабатываемая выборка содержит выбросы — наблюдения, плохо согласующиеся с остальными. В этих случаях пользуются робастными [5, 6], т. е. менее чувствительными к выбросам, чем МНК, методами оценивания. К их числу относится метод наименьших модулей (МНМ) [7], суть которого состоит в решении задачи:

$$J_1(\alpha) = \sum_{i=1}^n |\varepsilon_i| \rightarrow \min. \quad (3)$$

В монографии [8] показано, что задача (3) может быть сведена к задаче линейного программирования (ЛП):

$$\sum_{i=1}^n (u_i + v_i) \rightarrow \min, \quad (4)$$

$$\alpha_0 + \sum_{j=1}^m \alpha_j x_{ij} + u_i - v_i = y_i, \quad i = \overline{1, n}, \quad (5)$$

$$u_i \geq 0, \quad v_i \geq 0, \quad i = \overline{1, n}. \quad (6)$$

Противоположным по отношению к МНМ является метод антиробастного оценивания (МАО) [8], применяемый тогда, когда исследователь убежден в уникальности каждого наблюдения выборки и когда длина ее сильно ограничена, поэтому игнорировать какие-то наблюдения, используя робастные или близкие к ним методы, может оказаться нерациональным. Антиробастная оценка является решением задачи:

$$J_\infty(\alpha) = \max_k |\varepsilon_k| \rightarrow \min. \quad (7)$$

Такая оценка [8] определяется с помощью решения задачи ЛП, состоящей из ограничений (5), (6),

$$u_i + v_i \leq g, \quad i = \overline{1, n}, \quad (8)$$

и целевой функции

$$g \rightarrow \min. \quad (9)$$

Обобщением МНК, МНМ и МАО является так называемое L_ν -оценивание [9], состоящее в решении задачи:

$$J_\nu(\alpha) = \sum_{i=1}^n |\varepsilon_i|^\nu \rightarrow \min. \quad (10)$$

В [9] отмечено, что функция (10) при $\nu \geq 1$ выпукла вниз, а это означает, что при $\nu \geq 1$ существует единственный локальный минимум, который совпадает с глобальным. Как видно, при $\nu = 1$ L_ν -оценки совпадают с МНМ-оценками, при $\nu = 2$ — с МНК-оценками, а при $\nu \rightarrow \infty$ с МАО-оценками.

Для решения задачи (10) при $\nu \geq 1$ разработан метод, названный в [10] итеративным МНК, а в [11] — методом вариационно-взвешенных квадратичных приближений (МВВКП).

В работах [8, 12, 13] рассмотрен метод множественного оценивания (ММО) параметров линейной регрессии (1), в котором в качестве критериев минимизации используются два: J_1 и J_∞ , являющиеся полярными по отношению к выбросам. ММО предполагает решение двухкритериальной задачи ЛП с ограничениями (5), (6), (8) и с векторным критерием $J = (J_1, J_\infty)$. Ее решение можно получить, предварительно сформировав линейную свертку критериев J_1 и J_∞ с изменяющимися положительными коэффициентами λ :

$$J_\lambda = \lambda \cdot J_1 + (1 - \lambda) \cdot J_\infty, \quad \lambda \in [0, 1].$$

Тогда, равномерно разбивая отрезок $\lambda \in [0, 1]$ как можно большим количеством точек и решая для каждой из них задачу $\lambda \cdot \sum_{i=1}^n (u_i + v_i) + (1 - \lambda) \cdot g \rightarrow \min$ с ограничениями (5), (6) и (8), можно получить множество оценок ММО, представляющее собой множество Парето в критериальном пространстве (J_1, J_∞) . Для этого можно воспользоваться программным комплексом МОРМ [14].

Данная работа посвящена исследованию возможного симбиоза ММО с L_ν -оцениванием.

Множественное L_ν -оценивание

Для линейной регрессии (1) сформулируем задачу оценивания, в которой в качестве критериев минимизации выступают J_ν и J_∞ , $\nu \geq 1$. По смыслу такая задача ничем не отличается от формулировки ММО, а при $\nu = 1$ и вовсе эквивалентна ей. Однако можно заметить, что при $\nu \rightarrow \infty$ критерий $J_\nu \rightarrow J_\infty$, поэтому двухкритериальная задача с векторным критерием (J_ν, J_∞) при больших ν трансформируется в однокритериальную для МАО. Поэтому интерес вызывают только оценки, полученные предлагаемым методом при небольших значениях ν , которые могут быть лучше в некотором смысле, чем оценки, полученные известными на сегодняшний день методами.

Считая, что число ν задано, составим линейную свертку критериев J_ν и J_∞ :

$$J_{\lambda, \nu} = \lambda \cdot J_\nu + (1 - \lambda) \cdot J_\infty, \quad \lambda \in [0, 1].$$

Тогда, разбивая отрезок $\lambda \in [0, 1]$ некоторым количеством точек и решая для каждой из них оптимизационную задачу с целевой функцией

$$J_{\lambda, \nu}(\alpha) = \lambda \cdot \sum_{i=1}^n |\varepsilon_i|^\nu + (1 - \lambda) \cdot \max_k |\varepsilon_k| = \lambda \cdot \sum_{i=1}^n (u_i + v_i)^\nu + (1 - \lambda) \cdot g \rightarrow \min \quad (11)$$

и с ограничениями (5), (6) и (8), можно получить множество оценок линейной регрессии.

Решение задачи (11) с линейными ограничениями (5), (6) и (8) нетрудно получить при $\nu = 1$, когда она является задачей ЛП для любого λ , при $\lambda = 0$, когда она является линейной для любого ν , и при $\lambda = 1$, когда можно воспользоваться МВВКП. В остальных случаях она является задачей нелинейного программирования и может быть решена численно с использованием специализированных программ. Применение численных методов для решения задачи (11) с ограничениями (5), (6) и (8) сопряжено с двумя проблемами.

1. Неясно, как задавать начальные значения переменных, чтобы численный метод сходил к глобальному минимуму.

2. Функция (11) является неоднородной. Действительно, величина первого критерия $J_\nu = \sum_{i=1}^n |\varepsilon_i|^\nu$ зачастую может быть в значительной степени больше, чем величина второго $J_\infty = \max_k |\varepsilon_k|$. Поэтому равномерное разбиение отрезка $\lambda \in [0, 1]$ из-за доминирования первого критерия при минимизации может приводить к получению во всех точках, кроме $\lambda = 0$, одинаковых оценок.

Для решения этих проблем возведем критерий J_∞ в степень ν и рассмотрим задачу оценивания линейной регрессии по критериям J_ν и J_∞^ν , $\nu \geq 1$, которую будем называть множественным L_ν -оцениванием. Очевидно, что такая задача по смыслу эквивалентна задаче оценивания модели (1) по критериям J_ν и J_∞ . Для заданного ν линейная свертка критериев J_ν и J_∞^ν имеет вид:

$$J_{\lambda, \nu}^* = \lambda \cdot J_\nu + (1 - \lambda) \cdot J_\infty^\nu, \quad \lambda \in [0, 1].$$

Поскольку $J_\infty^\nu = \left(\max_k |\varepsilon_k| \right)^\nu = \max_k |\varepsilon_k|^\nu$, то после разбиения отрезка $\lambda \in [0, 1]$ точками для каждой из них нужно решить либо оптимизационную задачу с целевой функцией

$$J_{\lambda, \nu}^*(\alpha) = \lambda \cdot \sum_{i=1}^n (u_i + v_i)^\nu + (1 - \lambda) \cdot g^\nu \rightarrow \min \quad (12)$$

и с линейными ограничениями (5), (6), (8), либо с целевой функцией

$$J_{\lambda, \nu}^*(\alpha) = \lambda \cdot \sum_{i=1}^n (u_i + v_i)^\nu + (1 - \lambda) \cdot g^* \rightarrow \min, \quad (13)$$

с линейными ограничениями (5), (6) и нелинейными ограничениями

$$(u_i + v_i)^\nu \leq g^*, \quad i = \overline{1, n}. \quad (14)$$

Задача (12) с ограничениями (5), (6), (8) и задача (13) с ограничениями (5), (6), (14) являются эквивалентными. Вопрос о том, с какой из этих двух задач численные методы справляются эффективнее, пока остается открытым. Предпочтительнее, на наш взгляд, выглядит задача с целевой функцией (12), (5), (6), (8), поскольку область ее допустимых решений для любого ν представляет собой выпуклый многоугольник.

Очевидно, что функция (12) является однородной степени ν . Заметим, что все известные функции потерь — для МНМ, МНК, МАО, ММО и L_ν -оценивания — обладают свойством однородности. Функция (11) при $\nu > 1$ этим свойством не обладает.

Из-за однородности функции (12), во-первых, корректно разбивать отрезок $\lambda \in [0, 1]$ точками равномерно, во-вторых, для удобства множественного L_ν -оценивания позволительно нормировать исходные переменные $\bar{y}, x_1, x_2, \dots, x_m$ и переходить к оценкам стандартизованной линейной регрессии.

Проведем нормирование исходных переменных по правилам:

$$y_i^* = \frac{y_i - \bar{y}}{\sigma_y}, \quad x_{i1}^* = \frac{x_{i1} - \bar{x}_1}{\sigma_{x_1}}, \quad \dots, \quad x_{im}^* = \frac{x_{im} - \bar{x}_m}{\sigma_{x_m}}, \quad i = \overline{1, n},$$

где $\bar{y}, \bar{x}_1, \bar{x}_2, \dots, \bar{x}_m$ — средние значения переменных, а $\sigma_y, \sigma_{x_1}, \dots, \sigma_{x_m}$ — их среднеквадратические отклонения.

Тогда модель стандартизованной линейной регрессии имеет вид:

$$y_i^* = \beta_0 + \sum_{j=1}^m \beta_j x_{ij}^* + \xi_i, \quad i = \overline{1, n}, \quad (15)$$

где $\beta_0, \beta_1, \dots, \beta_m$ — неизвестные стандартизованные параметры; ξ_i — новая i -я ошибка аппроксимации.

Множественное L_ν -оценивание параметров стандартизованной регрессии (15) предполагает решение задачи:

$$J_{\lambda, \nu}^*(\beta) = \lambda \cdot \sum_{i=1}^n |\xi_i|^\nu + (1 - \lambda) \cdot \max_k |\xi_k|^\nu \rightarrow \min. \quad (16)$$

Нетрудно доказать, что оптимальное решение задачи (16) для заданного $\lambda \in [0,1]$ и $\nu \geq 1$ совпадает с оптимальным решением задачи $J_{\lambda,\nu}^*(\alpha) \rightarrow \min$. Для этого подставим в функцию (16) выраженные из (15) ошибки аппроксимации:

$$J_{\lambda,\nu}^*(\beta) = \lambda \cdot \sum_{i=1}^n \left| y_i^* - \beta_0 - \sum_{j=1}^m \beta_j x_{ij}^* \right|^\nu + (1 - \lambda) \cdot \max_k \left| y_i^* - \beta_0 - \sum_{j=1}^m \beta_j x_{ij}^* \right|^\nu.$$

С учетом правил нормирования переменных перепишем это выражение в виде:

$$\begin{aligned} J_{\lambda,\nu}^*(\beta) &= \lambda \cdot \sum_{i=1}^n \left| \frac{y_i - \bar{y}}{\sigma_y} - \beta_0 - \sum_{j=1}^m \beta_j \frac{x_{ij} - \bar{x}_j}{\sigma_{x_j}} \right|^\nu + \\ &+ (1 - \lambda) \cdot \max_k \left| \frac{y_i - \bar{y}}{\sigma_y} - \beta_0 - \sum_{j=1}^m \beta_j \frac{x_{ij} - \bar{x}_j}{\sigma_{x_j}} \right|^\nu = \\ &= \frac{1}{\sigma_y^\nu} \lambda \cdot \sum_{i=1}^n \left| y_i - \bar{y} - \beta_0 \sigma_y - \sum_{j=1}^m \beta_j \sigma_y \frac{x_{ij} - \bar{x}_j}{\sigma_{x_j}} \right|^\nu + \\ &+ \frac{1}{\sigma_y^\nu} (1 - \lambda) \cdot \max_k \left| y_i - \bar{y} - \beta_0 \sigma_y - \sum_{j=1}^m \beta_j \sigma_y \frac{x_{ij} - \bar{x}_j}{\sigma_{x_j}} \right|^\nu. \end{aligned}$$

Сделав в этом выражении замену переменных

$$\alpha_j = \beta_j \frac{\sigma_y}{\sigma_{x_j}}, \quad j = \overline{1, m}, \quad \alpha_0 = \bar{y} + \beta_0 \sigma_y - \sum_{j=1}^m \alpha_j \bar{x}_j, \quad (17)$$

получим $J_{\lambda,\nu}^*(\beta) = \frac{1}{\sigma_y^\nu} \left(\lambda \cdot \sum_{i=1}^n |\varepsilon_i|^\nu + (1 - \lambda) \cdot \max_k |\varepsilon_k|^\nu \right) = \frac{1}{\sigma_y^\nu} \cdot J_{\lambda,\nu}^*(\alpha)$. Отсюда следует, что минимизация функции $J_{\lambda,\nu}^*(\beta)$ равносильна минимизации функции $J_{\lambda,\nu}^*(\alpha)$. Формулы (17) устанавливают связь между параметрами стандартизованной (15) и исходной линейной регрессии (1).

Проводить множественное L_ν -оценивание можно, например, по следующему алгоритму.

1. Назначить область изменения параметра $\nu \in [1, \nu_{\text{наиб}}]$, где $\nu_{\text{наиб}}$ — его наибольшее значение. Выбрать $\nu_{\text{наиб}}$ можно следующим образом. Сначала оценить линейную регрессию по МАО и найти значение критерия $J_\infty(\bar{\alpha})$. Затем последовательно находить L_ν -оценки модели по МВВКП, последовательно увеличивая параметр ν от единицы до тех пор, пока критерий антиробастности регрессии $\max_k |\varepsilon_k|$ будет достаточно отличен от $J_\infty(\bar{\alpha})$. Как отмечено выше, проводить множественное L_ν -оценивание при больших ν вряд ли имеет смысл.

2. Провести нормирование переменных.

3. Равномерно разбить отрезки $\nu \in [1, \nu_{\text{наиб}}]$ и $\lambda \in [0,1]$ точками и в каждой такой точке для нормированных данных решить либо задачу (12) с ограничениями (5), (6), (8), либо задачу (13) с ограничениями (5), (6), (14). Поскольку переменные нормированы, то в качестве их начальных приближений при реализации численных методов можно указать нули. Для большей надежности при решении задачи нелинейного программирования для некоторого узла сети (ν, λ) желательно использовать в качестве начальных приближений оценки соседних узлов.

4. По формулам (17) для каждого узла сети (ν, λ) от оценок стандартизованной регрессии перейти к оценкам исходной регрессии (1).

Выбор из полученного множества оценок с наилучшими свойствами пока остается открытой научной задачей.

Моделирование

Для демонстрации метода множественного L_ν -оценивания решалась задача моделирования пассажирских железнодорожных перевозок в Иркутской области. Для этого были использованы статистические данные (табл. 1) за 2000–2020 гг. по следующим переменным:

y — отправление пассажиров железнодорожным транспортом общего пользования (тыс. человек);

x_1 — число собственных легковых автомобилей на 1000 человек населения (штук);

x_2 — средние потребительские цены на полет в салоне экономического класса самолета (в расчете на 1000 км пути) за декабрь (рублей).

Таблица 1

Статистические данные

Год	y	x_1	x_2	Год	y	x_1	x_2
2000	22133	138,8	1000,63	2011	18059	224,3	5863,89
2001	23216	140	1058,38	2012	17266	251,5	5888,33
2002	23363	147,2	1272,93	2013	14930	271,8	3554,45
2003	25122	151,6	1804,49	2014	14161	270,5	3641,52
2004	26435	156,5	1962,69	2015	12967	271,3	5095,42
2005	26247	143,3	2819,53	2016	12998	242,7	5297,99
2006	27817	154,8	3529,62	2017	12003	246,2	5129,7
2007	28501	169,2	4811,21	2018	11796	245,6	4669,4
2008	28757	188,2	6590,97	2019	11907	254,9	4416,62
2009	22427	189,8	6202,74	2020	9294	261,6	4653,39
2010	19774	202,6	5633,74				

Значения параметра ν задавались равными 1, 1,5, 2, 2,5, 3, 3,5, 4, 4,5, 5, 10 и 20. Отрезок $\lambda \in [0,1]$ был разбит равномерно девятью точками. Сначала данные были нормированы, а затем для каждого узла из сети (ν, λ) с помощью онлайн-решателя APMonitor решалась задача (12) с ограничениями (5), (6), (8). В качестве начальных приближений переменных были указаны нули. Результаты моделирования представлены в табл. 2.

В каждой ячейке табл. 2 приведен вектор $A = (\bar{\alpha}_0, \bar{\alpha}_1, \bar{\alpha}_2, J_\nu(\bar{\alpha}), J_\infty^\nu(\bar{\alpha}), J_1(\bar{\alpha}), J_\infty(\bar{\alpha}))$, где $\bar{\alpha}_0, \bar{\alpha}_1, \bar{\alpha}_2$ — оценки линейной регрессии, пересчитанные по формулам (17), $J_\nu(\bar{\alpha}), J_\infty^\nu(\bar{\alpha}), J_1(\bar{\alpha}), J_\infty(\bar{\alpha})$ — значения функций потерь для оценок $\bar{\alpha}_0, \bar{\alpha}_1, \bar{\alpha}_2$.

По табл. 2 видно, что в критериальном пространстве (J_ν, J_∞^ν) при $\nu = 1$ получено 4 паретовских вершины, а начиная с $\nu = 1.5$ паретовская вершина получается для каждого значения λ , т. е. на отрезке $\lambda \in [0,1]$ образуется выпуклая вниз фигура. Более того, в критериальном пространстве (J_1, J_∞) можно наблюдать точно такую же картину (рис.).

По рис. можно проследить, как быстро с увеличением ν двухкритериальная задача с векторным критерием (J_ν, J_∞^ν) трансформируется в однокритериальную задачу антиробастного оценивания.

Затем методом «идеальной» точки из табл. 1 были выбраны оптимальные оценки с точки зрения критериев J_1 и J_∞ . Сначала эти критерии были нормированы по правилам

$$\bar{J}_1(\lambda, \nu) = \frac{J_1(\lambda, \nu) - \min J_1}{\max J_1 - \min J_1}, \quad \bar{J}_\infty(\lambda, \nu) = \frac{J_\infty(\lambda, \nu) - \min J_\infty}{\max J_\infty - \min J_\infty}.$$

После чего была назначена «идеальная» точка (0,0) и выбрана альтернатива, наиболее близкая к этой точке с точки зрения евклидова расстояния

$$r = \min_{\lambda, \nu} \sqrt{(\bar{J}_1(\lambda, \nu))^2 + (\bar{J}_\infty(\lambda, \nu))^2}.$$

Ею оказалась альтернатива, для которой $\nu = 1,5, \lambda = 0,7$. Уравнение линейной регрессии для этих параметров имеет вид:

$$\bar{y} = 41847 - 128,95x_1 + 1,015x_2. \quad (18)$$

Критерии адекватности модели (18) $J_1 = 48490, J_\infty = 4525$.

Уравнение (18) удовлетворяет содержательному смыслу решаемой задачи. Переменная x_1 оказывает обратное влияние на y , поскольку рост благосостояния населения, а значит, и числа легковых

Таблица 2

Результаты моделирования

λ v	1	1,5	2	2,5	3	3,5	4	4,5	5	10	20
0	(43522; -131,95; 0,921; 50903; 3997; 50903; 3997)	(43522; -131,95; 0,921; 2,75E+06; 2,53E+05; 50903; 3997)	(43522; -131,95; 0,921; 1,54E+08; 1,60E+07; 50903; 3997)	(43522; -131,95; 0,921; 8,80E+09; 1,01E+09; 50903; 3997)	(43522; -131,95; 0,921; 5,14E+11; 6,39E+10; 50903; 3997)	(43522; -131,95; 0,921; 3,04E+13; 4,04E+12; 50903; 3997)	(43522; -131,95; 0,921; 1,82E+15; 2,55E+14; 50903; 3997)	(43522; -131,95; 0,921; 1,10E+17; 1,61E+16; 50903; 3997)	(43522; -131,95; 0,921; 6,66E+18; 1,02E+18; 50903; 3997)	(43522; -131,95; 0,921; 5,44E+36; 1,04E+36; 50903; 3997)	(43522; -131,95; 0,921; 4,84E+72; 1,08E+72; 50903; 3997)
0,1	(41216; -125,87; 1,088; 49079; 4058; 49079; 4058)	(41716; -127,19; 1,052; 2,65E+06; 2,57E+05; 49449; 4045)	(42139; -128,3; 1,021; 1,48E+08; 1,63E+07; 49762; 4034)	(42350; -128,86; 1,006; 8,45E+09; 1,03E+09; 49918; 4028)	(42490; -129,23; 0,996; 4,91E+11; 6,52E+10; 50021; 4024)	(42595; -129,50; 0,988; 2,90E+13; 4,13E+12; 50098; 4022)	(42676; -129,72; 0,982; 1,73E+15; 2,61E+14; 50158; 4020)	(42744; -129,90; 0,977; 1,04E+17; 1,65E+16; 50208; 4018)	(42802; -130,05; 0,973; 6,31E+18; 1,04E+18; 50252; 4016)	(43135; -130,92; 0,949; 5,08E+36; 1,07E+36; 50498; 4007)	(43337; -131,45; 0,934; 4,48E+72; 1,11E+72; 50647; 4002)
0,2	(41216; -125,87; 1,088; 49079; 4058; 49079; 4058)	(41399; -126,35; 1,075; 2,65E+06; 2,58E+05; 49214; 4053)	(41821; -127,47; 1,044; 1,47E+08; 1,63E+07; 49526; 4042)	(42019; -127,99; 1,030; 8,41E+09; 1,04E+09; 49673; 4037)	(42144; -128,31; 1,021; 4,89E+11; 6,56E+10; 49766; 4033)	(42235; -128,56; 1,014; 2,88E+13; 4,16E+12; 49832; 4031)	(42304; -128,73; 1,009; 1,71E+15; 2,64E+14; 49884; 4029)	(42364; -128,90; 1,005; 1,03E+17; 1,67E+16; 49928; 4028)	(42417; -129,04; 1,007; 6,22E+18; 1,06E+18; 49967; 4026)	(42812; -130,07; 0,972; 4,93E+36; 1,09E+36; 50259; 4016)	(43145; -130,95; 0,948; 4,30E+72; 1,14E+72; 50506; 4007)
0,3	(41216; -125,87; 1,088; 49079; 4058; 49079; 4058)	(41308; -126,11; 1,081; 2,65E+06; 2,58E+05; 49147; 4055)	(41715; -127,19; 1,052; 1,47E+08; 1,64E+07; 49448; 4045)	(41908; -127,69; 1,038; 8,40E+09; 1,04E+09; 49591; 4040)	(42028; -128,01; 1,029; 4,88E+11; 6,58E+10; 49679; 4037)	(42110; -128,23; 1,023; 2,87E+13; 4,17E+12; 49740; 4034)	(42173; -128,39; 1,019; 1,71E+15; 2,64E+14; 49787; 4033)	(42226; -128,53; 1,015; 1,03E+17; 1,68E+16; 49826; 4031)	(42273; -128,66; 1,012; 6,20E+18; 1,06E+18; 49861; 4030)	(42666; -129,69; 0,983; 4,90E+36; 1,10E+36; 50152; 4020)	(43075; -130,91; 0,963; 4,17E+72; 1,18E+72; 50478; 4015)
0,4	(41216; -125,87; 1,088; 49079; 4058; 49079; 4058)	(41354; -126,43; 1,073; 2,63E+06; 2,65E+05; 49015; 4126)	(41672; -127,10; 1,054; 1,47E+08; 1,64E+07; 49394; 4056)	(41851; -127,54; 1,042; 8,40E+09; 1,04E+09; 49549; 4041)	(41966; -127,85; 1,034; 4,88E+11; 6,58E+10; 49634; 4038)	(42047; -128,06; 1,028; 2,87E+13; 4,18E+12; 49693; 4036)	(42105; -128,21; 1,024; 1,71E+15; 2,65E+14; 49737; 4034)	(42154; -128,34; 1,020; 1,03E+17; 1,68E+16; 49773; 4033)	(42197; -128,45; 1,017; 6,20E+18; 1,07E+18; 49805; 4032)	(42597; -129,71; 1,001; 4,85E+36; 1,13E+36; 50134; 4030)	(43034; -130,92; 0,948; 4,10E+72; 1,22E+72; 50468; 4021)
0,5	(42966; -133,79; 0,87; 47531; 5236; 47531; 5236)	(41574; -127,49; 1,044; 2,61E+06; 2,82E+05; 48763; 4299)	(41782; -127,73; 1,037; 1,46E+08; 1,75E+07; 49181; 4180)	(41963; -127,86; 1,032; 8,35E+09; 1,08E+09; 49423; 4109)	(42088; -129,05; 1,049; 4,81E+11; 7,36E+10; 49327; 4190)	(42126; -129,10; 1,051; 2,84E+13; 4,63E+12; 49616; 4157)	(42151; -129,10; 1,052; 1,69E+15; 2,91E+14; 49694; 4131)	(42173; -129,10; 1,051; 1,01E+17; 1,83E+16; 49752; 4112)	(42195; -129,10; 1,049; 6,13E+18; 1,15E+18; 49798; 4096)	(42515; -129,71; 1,021; 4,81E+36; 1,16E+36; 50111; 4041)	(42980; -130,92; 0,987; 4,05E+72; 1,26E+72; 50453; 4028)
0,6	(43755; -145,35; 1,178; 46949; 6495; 46949; 6495)	(41847; -128,95; 1,015; 2,60E+06; 3,29E+05; 48364; 4770)	(42061; -130,15; 1,064; 1,44E+08; 1,99E+07; 48989; 4460)	(42186; -130,48; 1,075; 8,23E+09; 1,28E+09; 49295; 4389)	(42227; -130,40; 1,076; 4,77E+11; 8,08E+10; 49473; 4323)	(42242; -130,25; 1,074; 2,81E+13; 5,09E+12; 49566; 4271)	(42251; -130,11; 1,072; 1,67E+15; 3,21E+14; 49669; 4232)	(42260; -130,00; 1,070; 1,00E+17; 2,02E+16; 49728; 4202)	(42269; -129,90; 1,066; 6,06E+18; 1,27E+18; 49774; 4177)	(42514; -129,90; 1,029; 4,78E+36; 1,23E+36; 50098; 4063)	(42962; -130,92; 0,991; 4,04E+72; 1,28E+72; 50448; 4030)
0,7	(43755; -145,35; 1,178; 46949; 6495; 46949; 6495)	(42051; -131,29; 1,067; 2,59E+06; 3,29E+05; 48364; 4770)	(42235; -131,85; 1,099; 1,44E+08; 2,14E+07; 48959; 4625)	(42322; -131,81; 1,101; 8,19E+09; 1,37E+09; 49269; 4519)	(42341; -131,52; 1,098; 4,75E+11; 8,71E+10; 49449; 4433)	(42340; -131,21; 1,094; 2,80E+13; 5,50E+12; 49566; 4366)	(42336; -130,96; 1,090; 1,66E+15; 3,47E+14; 49648; 4316)	(42332; -130,75; 1,086; 9,98E+16; 2,19E+16; 49708; 4277)	(42333; -130,58; 1,081; 6,03E+18; 1,38E+18; 49755; 4246)	(42535; -130,23; 1,038; 4,74E+36; 1,34E+36; 50080; 4099)	(42949; -130,94; 0,995; 4,04E+72; 1,29E+72; 50439; 4033)
0,8	(43755; -145,35; 1,178; 46949; 6495; 46949; 6495)	(42199; -133,13; 1,110; 2,59E+06; 3,50E+05; 48273; 4968)	(42378; -133,25; 1,128; 1,44E+08; 2,27E+07; 48933; 4762)	(42436; -132,92; 1,124; 8,18E+09; 1,46E+09; 49246; 4628)	(42436; -132,46; 1,117; 4,74E+11; 9,27E+10; 49428; 4525)	(42421; -132,02; 1,110; 2,79E+13; 5,86E+12; 49547; 4447)	(42405; -131,67; 1,105; 1,66E+15; 3,70E+14; 49629; 4387)	(42394; -131,39; 1,099; 9,95E+16; 2,34E+16; 49691; 4341)	(42388; -131,16; 1,094; 6,01E+18; 1,48E+18; 49738; 4304)	(42557; -130,53; 1,046; 4,72E+36; 1,45E+36; 50065; 4132)	(42944; -130,97; 0,997; 4,03E+72; 1,33E+72; 50439; 4039)
1	(43755; -145,35; 1,178; 46949; 6495; 46949; 6495)	(42346; -134,76; 1,146; 2,58E+06; 3,68E+05; 48214; 5137)	(42497; -134,43; 1,152; 1,44E+08; 2,38E+07; 48911; 4877)	(42534; -133,87; 1,143; 8,17E+09; 1,53E+09; 49227; 4721)	(42517; -133,25; 1,133; 4,74E+11; 9,76E+10; 49410; 4604)	(42491; -132,72; 1,124; 2,79E+13; 6,19E+12; 49530; 4516)	(42466; -132,29; 1,117; 1,66E+15; 3,92E+14; 49614; 4449)	(42448; -131,95; 1,111; 9,94E+16; 2,48E+16; 49675; 4397)	(42438; -131,68; 1,105; 6,00E+18; 1,57E+18; 49725; 4356)	(42573; -130,78; 1,053; 4,72E+36; 1,55E+36; 50051; 4160)	(42951; -131,10; 1,001; 4,03E+72; 1,44E+72; 50430; 4054)

автомобилей, снижает спрос на железнодорожные перевозки. Переменная x_2 оказывает прямое влияние на y : чем выше цены на полеты в самолетах, тем ниже спрос на авиаперевозки, и тем выше спрос на железнодорожные перевозки.

Заключение

В работе предложен метод множественного L_p -оценивания неизвестных параметров линейной регрессии. В результате оценивания формируется множество различных оценок линейной регрессии. Вопросы выбора из этого множества оценок с наилучшими свойствами будут исследованы в дальнейших работах авторов.

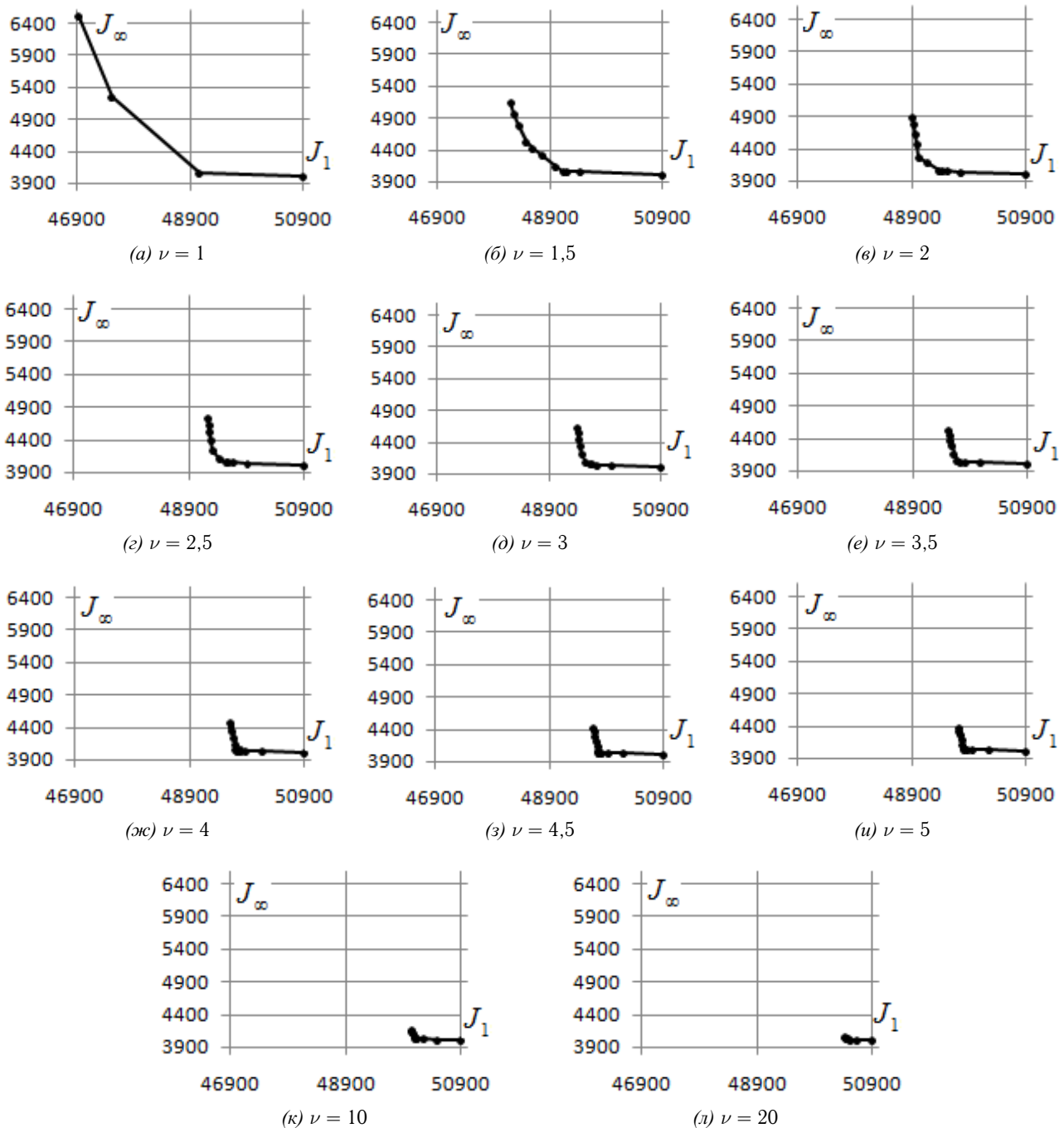


Рис. Вершины Парето для разных ν в критериальном пространстве (J_1, J_∞)

ЛИТЕРАТУРА

1. Brook R. J., Arnold G. C. *Applied Regression Analysis and Experimental Design*. CRC Press; 2018.
2. Lawrence K. D. *Robust Regression: Analysis and Applications*. Routledge; 2019.
3. Mahanty C., Kumar R., Mishra B. K. Analyses the Effects of COVID-19 Outbreak on Human Sexual Behaviour Using Ordinary Least-Squares Based Multivariate Logistic Regression. *Quality & Quantity*. 2021;55(4):1239–1259. Режим доступа: <https://link.springer.com/article/10.1007/s11135-020-01057-8>.
4. Aloisio A., Alaggio R., Fragiaco M. Dynamic Identification of a Masonry Façade from Seismic Response Data Based on an Elementary Ordinary Least Squares Approach. *Engineering Structures*. 2019;197:109415. Режим доступа: <https://www.sciencedirect.com/science/article/abs/pii/S0141029618322697>.
5. Huber P. J. Robust Estimation of a Location Parameter. *Breakthroughs in Statistics*. 1992:492–518.
6. Wilcox R. R. *Introduction to Robust Estimation and Hypothesis Testing*. Academic Press; 2011.

7. Bloomfield P., Steiger W. L. *Least Absolute Deviations: Theory, Applications, and Algorithms*. Boston: Birkhäuser; 1983.
8. Носков С. И. *Технология моделирования объектов с нестабильным функционированием и неопределенностью в данных*. Иркутск: Облформпечать; 1996.
9. Демиденко Е. З. *Линейная и нелинейная регрессии*. М.: Финансы и статистика; 1981.
10. Fletcher R., Grant J. A., Hebden H. D. The Calculation of Linear Least L_p -approximations. *Computer Journal*. 1971;14(3).
11. Мудров В. И., Кушко В. Л. *Методы обработки измерений: Квазиправдоподобные оценки*. М.: Радио и связь; 1983.
12. Носков С. И., Базилевский М. П. Множественное оценивание параметров и критерий согласованности поведения в регрессионном анализе. *Вестник Иркутского государственного технического университета*. 2018;4(135):101–110.
13. Носков С. И., Баенхаева А. В. Множественное оценивание параметров линейного регрессионного уравнения. *Современные технологии. Системный анализ. Моделирование*. 2016;3(51):133–138.
14. Баенхаева А. В., Базилевский М. П., Носков С. И. Программный комплекс множественного оценивания регрессионных моделей. *Информационные технологии и проблемы математического моделирования сложных систем*. 2016;17:38–44.