УДК 519.25
DOI: 10.26456/2219-1453/2024.4.231–241

СГЛАЖИВАНИЕ СТАТИСТИЧЕСКИХ ДАННЫХ ЭКОНОМИЧЕСКИХ ПОКАЗАТЕЛЕЙ С ИСПОЛЬЗОВАНИЕМ КОЭФФИЦИЕНТОВ ИХ ЗНАЧИМОСТИ

Г.М. Соломаха, С.В. Хижняк

ФГБОУ ВО «Тверской государственный университет», г. Тверь

Статья посвящена исследованию вопросов, связанных со сглаживанием статистических данных экономических показателей. Целью данной работы является уточнение модели расчета параметров линейной и нелинейной регрессии экономических показателей в случае, когда эти данные имеют разную значимость. Для этого предлагается использовать коэффициенты значимости для статистических данных. При этом для нахождения параметров регрессии возникает необходимость решения нелинейных оптимизационных задач. Их решение представлено в статье. Приводится пример использования предложенного подхода для анализа производства автомобилей в странах для временного периода, включающего годы пандемии. Научная новизна полученных результатов заключается в том, что приведенные в статье модели в отличие от существующих моделей сглаживания данных позволяют учесть разную значимость статистических данных при их анализе и прогнозировании экономических показателей.

Ключевые слова: *линейное и нелинейное сглаживание данных, метод наименьших квадратов, сглаживание статистических данных.*

При исследовании экономических показателей широко используются их статистические данные. Действительно, анализ статистических данных за предшествующие периоды позволяет получить их аппроксимирующие зависимости, а также спрогнозировать величину соответствующего экономического показателя в будущие периоды. Это, в свою очередь, помогает планированию производства и развитию бизнеса в целом.

В настоящее время имеется большое количество работ, в которых рассматриваются анализ и прогнозирование рядов экономических показателей, в частности [1, 2]. Следует отметить, что данная задача может быть рассмотрена в общем случае как задача идентификации [3] сложных динамических систем, однако для ее решения требуется ряд дополнительных параметров, что затрудняет использование этого подхода для рассматриваемой задачи. Кроме того, рассматриваемая задача может быть исследована как задача исследования операций [4], однако, обычно рассматривать эту задачу в условиях неопределенности и риска не представляется возможным, поскольку пользователи статистических данных не обладают информацией о том, какими способами и с какой погрешностью получены эти статистические данные в публикуемых документах. Наиболее широко применяется для анализа статистических данных экономических показателей метод наименьших квадратов. Суть его состоит в выборе модели изменения исследуемого показателя (линейной или нелинейной) с последующим нахождением параметров этой модели, исходя из решения оптимизационной задачи нахождения минимума суммы квадратов отклонений имеющихся статистических данных от соответствующих значений указанной гипотетической функции. Тем самым негласно считается, что все статистические данные в рассматриваемом исследовании экономических показателей равнозначны, то есть имеют одинаковую значимость. Однако это часто бывает неверно. Так, некоторые показатели могут вызывать сомнения у их пользователя. Действительно, ведь встречаются и технические ошибки и опечатки. Кроме того, например, в условиях пандемии в 2020 и 2021 гг. значения многих экономических показателей резко изменились. Поэтому использовать их по той же схеме, что и в предыдущие годы, представляется невозможным. Нельзя и просто исключить статистические данные за годы пандемии, поскольку экономические системы продолжали функционировать и в этот временной период. Поэтому возникает необходимость разработки механизма для учета статистических данных за этот период. При этом желательно, чтобы предложенный подход не был направлен на рассмотрение частных случаев (сомнительных данных или пандемии, или дезинформации, или других), а был наиболее общим. То есть этот подход призван обеспечить решение задач аппроксимации и прогнозирования экономических показателей в указанных случаях или в других ситуациях как частные решения общего механизма для этого подхода. Ясно, что такой подход может служить основой для решения ряда других задач, например, каким образом учитывать статистические данные за годы пандемии в исследованиях для будущих временных периодов.

В соответствии с этим будем рассматривать статистические данные некоторого экономического показателя вместе с весовыми коэффициентами, которые характеризуют значимость этого показателя. Очевидно, если взять все эти весовые коэффициенты равными 1, то получим стандартный метод наименьших квадратов. Если же некоторое статистическое данное исключаем из рассмотрения, то его весовой коэффициент (значимость) равен 0. При этом весовые коэффициенты всех статистических данных в каждом исследовании могут изменяться от 0 до 1 включительно. Тогда, если у исследователя возникают сомнения в некоторых данных, то он может уменьшить их значимость путем уменьшения соответствующих весовых коэффициентов для них. Понятно, что выбор этого коэффициента представляет собой отдельную задачу, которая является оптимизационной задачей и может быть решена на базе предложенного подхода. В ряде случаев, когда имеется несколько временных периодов, для которых значимость статистических данных ниже 1, например, в период пандемии, то можно оценить коэффициенты значимости данных в эти периоды, считая их одинаковыми. Тогда эти коэффициенты также представляется возможным найти путем решения соответствующей оптимизационной задачи на базе предлагаемого общего подхода. Следует подчеркнуть, что предлагаемые коэффициенты значимости характеризуют их вес в целевой функции в задаче оптимизации, а не общую значимость (важность) этого показателя.

Целью работы является уточнение модели расчета параметров линейной и нелинейной регрессии экономических показателей для случая, когда эти данные имеют разную значимость.

Для достижения данной цели, прежде всего, остановимся на линейной модели аппроксимации статистических данных. Пусть имеется n значений переменной x, именно $x_1, x_2, \ldots, x_n$, а также n соответствующих значений y, то есть $y_1, y_2, \ldots, y_n$. Например, x - это временной период (год), а y - объем выпускаемой продукции предприятием в соответствующий период. Рассмотрим задачу сглаживания зависимости между переменными x и y. Такую сглаженную зависимость обозначим $y = f(x)$. Вид этой функции может быть выбран различным. Сначала рассмотрим случай линейной функции $y(x) = ax + b$, где a и b - параметры модели, которые необходимо определить. Графиком этой функции служит прямая, которая называется линией тренда. Для вычисления оптимальных значений коэффициентов a и b линейной модели широко используется метод наименьших квадратов (см., например [1,2]). Тогда a и b выбираются так, чтобы сумма квадратов отклонений значений, найденных по выражению $y = f(x)$ в указанных точках, от фактических значений y_i была минимальной для рассматриваемых n пар значений (x_i, y_i). Таким образом, должно выполняться условие

$$S(a,b) = \sum_{i=1}^{n} (ax_i + b - y_i)^2 \to \min_{a,b} . \qquad (1)$$

Однако в стандартной линейной модели, которой соответствует задача (1), все значения данных в ней считаются одинаково значимыми, что может не соответствовать реальной ситуации. Чтобы избежать этого в формулу целесообразно добавить весовые коэффициенты V_i, с помощью которых можно задать значимость каждого значения y_i в диапазоне от 0 до 1. При $V_i = 1$ значение y_i в (1) имеет наибольшую значимость, при $V_i = 0$ значение y_i считается незначимым, а с помощью промежуточных значений можно определять степень значимости данных, например, исходя из анализа событий года, в который было получено данное конкретное значение. При добавлении весовых коэффициентов V_i для данных n пар значений (x_i, y_i) приходим к задаче оптимизации

$$S(a,b) = \sum_{i=1}^{n} V_i \cdot (ax_i + b - y_i)^2 \to \min_{a,b} . \qquad (2)$$

При этом в задаче (2) минимизация осуществляется по двум неизвестным: a и b. Для нахождения минимума в (2) находим частные производные по неизвестным a и b, а затем приравниваем их к нулю. В результате получим систему двух линейных уравнений с двумя неизвестными

$$\begin{cases} \left(\sum_{i=1}^{n} V_i x_i^2\right) \cdot a + \left(\sum_{i=1}^{n} V_i x_i\right) \cdot b = \sum_{i=1}^{n} V_i x_i y_i \\ \left(\sum_{i=1}^{n} V_i x_i\right) \cdot a + \left(\sum_{i=1}^{n} V_i\right) \cdot b = \sum_{i=1}^{n} V_i y_i \end{cases} \quad (3)$$

В результате решения системы (3) найдем оптимальные значения параметров a и b линейной зависимости. Соответственно, в отличие от стандартного метода наименьших квадратов, в формулах (4) для их вычисления появятся V_i - весовые коэффициенты значимости:

$$\begin{cases} a = \dfrac{\left(\sum_{i=1}^{n} V_i x_i y_i\right) \cdot \left(\sum_{i=1}^{n} V_i\right) - \left(\sum_{i=1}^{n} V_i y_i\right) \cdot \left(\sum_{i=1}^{n} V_i x_i\right)}{\left(\sum_{i=1}^{n} V_i x_i^2\right) \cdot \left(\sum_{i=1}^{n} V_i\right) - \left(\sum_{i=1}^{n} V_i x_i\right)^2} \\ b = \dfrac{\left(\sum_{i=1}^{n} V_i y_i\right) \cdot \left(\sum_{i=1}^{n} V_i x_i^2\right) - \left(\sum_{i=1}^{n} V_i x_i y_i\right) \cdot \left(\sum_{i=1}^{n} V_i x_i\right)}{\left(\sum_{i=1}^{n} V_i x_i^2\right) \cdot \left(\sum_{i=1}^{n} V_i\right) - \left(\sum_{i=1}^{n} V_i x_i\right)^2} \end{cases} \quad (4)$$

Чтобы убедиться, что найденные из решения системы уравнений (3) значения a и b обеспечивают минимум целевой функции в (2), проверим достаточные условия [5] минимума целевой функции, состоящие в том, что все угловые миноры матрицы Гессе частных производных второго порядка положительны. Для этого сначала найдем частные производные второго порядка этой целевой функции:

$$S''_{aa}(a,b) = 2\sum_{i=1}^{n} V_i x_i^2; \quad S''_{ab}(a,b) = S''_{ab}(a,b) = 2\sum_{i=1}^{n} V_i x_i;$$

$$S''_{bb}(a,b) = 2\sum_{i=1}^{n} V_i.$$

Тогда угловой минор первого порядка матрицы Гессе

$$\Delta_1 = S''_{aa}(a,b) = 2\sum_{i=1}^{n} V_i x_i^2 > 0\,.$$

Найдем угловой минор второго порядка матрицы Гессе целевой функции:

$$\Delta_2 = S''_{aa}(a,b) \cdot S''_{bb}(a,b) - \left(S''_{ab}(a,b)\right)^2 = 4\left(\sum_{i=1}^{n} V_i\right) \cdot \sum_{i=1}^{n} V_i x_i^2 - 4\left(\sum_{i=1}^{n} V_i x_i\right)^2 =$$

$$= 4\left(\sum_{i=1}^{n} V_i^2 x_i^2 + \sum_{i=1}^{n} \sum_{j=1(j\neq i)}^{n} V_i V_j x_i^2 - \sum_{i=1}^{n} V_i^2 x_i^2 - \sum_{i=1}^{n} \sum_{j>i}^{n} 2V_i V_j x_i x_j \right).$$

В итоге, получаем, что

$$\Delta_2 = 4\sum_{i=1}^{n} \sum_{j>i}^{n} V_i V_j \left(x_i^2 - 2x_i x_j + x_j^2 \right) = 4\sum_{i=1}^{n} \sum_{j>i}^{n} V_i V_j (x_i - x_j)^2 > 0 .$$

Таким образом, достаточные условия минимума функции двух переменных выполнены, следовательно, a и b, определяемые выражениями (4), обеспечивают минимум целевой функции (2). Отметим, что поскольку a и b в (4) зависят от значений коэффициентов значимости статистических данных, то есть $a = a(\bar{V})$ и $b = b(\bar{V})$, где $\bar{V} = (V_1, V_2, ..., V_n)$, то уравнение прямой, сглаживающей эти данные, имеет следующий вид:

$$y(x,\bar{V}) = f(x,\bar{V}) = a(\bar{V}) \cdot x + b(\bar{V}). \tag{5}$$

В частном случае анализа экономических показателей во времени в качестве x_i выступают временные промежутки (чаще всего годы или месяцы), а в качестве y_i берутся соответствующие статистические данные рассматриваемого экономического показателя (например, объема выпускаемой продукции предприятием).

Теперь перейдем к рассмотрению нелинейной квадратичной модели (модели квадратичного тренда). Формула квадратичной модели записывается как $y = ax^2 + bx + c$, где a, b и c - это константы, причем графическим представлением этой модели является непосредственно график параболы. Теперь целевая функция с учетом весовых коэффициентов примет следующий вид:

$$S(a,b,c) = \sum_{i=1}^{n} V_i \left(ax_i^2 + bx_i + c - y_i \right)^2 \to \min_{a,b,c}, \tag{6}$$

где a, b и c неизвестные, подлежащие нахождению в результате решения оптимизационной задачи (6). Для их определения (в соответствии с необходимым условием экстремума функции трех переменных) запишем выражения для частных производных целевой функции в (6) по этим переменным и приравняем их к нулю. Получаем

$$\begin{cases} S'_a(a,b,c) = 0 \\ S'_b(a,b,c) = 0 \\ S'_c(a,b,c) = 0 \end{cases} \Rightarrow \begin{cases} 2\sum_{i=1}^{n} V_i \left(ax_i^2 + bx_i + c - y_i \right) x_i^2 = 0 \\ 2\sum_{i=1}^{n} V_i \left(ax_i^2 + bx_i + c - y_i \right) x_i = 0 \\ 2\sum_{i=1}^{n} V_i (ax_i^2 + bx_i + c - y_i) = 0. \end{cases}$$

После элементарных преобразований полученной системы трех уравнений с тремя неизвестными приходим к необходимости решения системы

$$\begin{cases} \left(\sum_{i=1}^{n} V_i x_i^4\right) \cdot a + \left(\sum_{i=1}^{n} V_i x_i^3\right) \cdot b + \left(\sum_{i=1}^{n} V_i x_i^2\right) \cdot c = \sum_{i=1}^{n} V_i x_i^2 y_i \\ \left(\sum_{i=1}^{n} V_i x_i^3\right) \cdot a + \left(\sum_{i=1}^{n} V_i x_i^2\right) \cdot b + \left(\sum_{i=1}^{n} V_i x_i\right) \cdot c = \sum_{i=1}^{n} V_i x_i y_i \\ \left(\sum_{i=1}^{n} V_i x_i^2\right) \cdot a + \left(\sum_{i=1}^{n} V_i x_i\right) \cdot b + \left(\sum_{i=1}^{n} V_i\right) \cdot c = \sum_{i=1}^{n} V_i y_i \end{cases} \tag{7}$$

относительно неизвестных a, b и c. При этом в (7) V_i представляют собой весовые коэффициенты значимости соответствующих статистических данных. В результате решения системы (7) одним из стандартных подходов, например методами Крамера или Гаусса, находится набор значений a, b и c, обеспечивающий достижение минимума в задаче (6). При этом аналогично линейному случаю можно показать, что для этого набора a, b, c выполнено достаточное условие минимума в (6), состоящее в положительности всех угловых миноров для матрицы Гессе вторых частных производных целевой функции в (6). При анализе во времени экономических показателей на основе рассмотренной нелинейной модели (как и в случае линейной модели) в качестве x_i выступают временные промежутки, а в качестве y_i берутся соответствующие статистические данные рассматриваемого экономического показателя.

Следует отметить, что данный подход с введением коэффициентов значимости данных можно использовать и в общем случае наличия корреляционной таблицы для двух случайных величин X и Y. Тогда каждому набору значений (x, y) также соответствует свой коэффициент значимости. Это позволяет, например, при расчете выборочного коэффициента корреляции учесть разную значимость указанных наборов значений. Также может быть рассмотрена, и нелинейная квадратичная модель зависимости случайных величин и рассчитана силы их связи с учетом значений коэффициентов значимости для пар значений (x, y).

Необходимо подчеркнуть, что к выбору коэффициентов значимости в выражениях (2), (6) или аналогичных им надо подходить ответственно, не следует безосновательно делать их значения меньшими, чем 1. Однако, если есть информация об аномальных периодах для исследуемых экономических процессов, то может быть целесообразно установить для них коэффициент значимости статистических данных ниже единицы. При этом коэффициент значимости равный нулю соответствует полному исключению из анализа данных указанного временного периода. Вместе с тем использование коэффициентов значимости статистических данных дает возможность исследования нестандартных случаев и, как следствие, осуществлять прогноз значений исследуемых экономических показателей более обоснованно. Кроме того, коэффициенты значимости для отдельных «подозрительных» данных в некоторые временные периоды могут быть оценены, например, по реальным статистическим данным для других последующих периодов.

Рассмотрим примеры нахождения индексов значимости для статистических данных мирового автомобилестроения. В табл. 1 представлены объемы производства автомобилей в трех странах (Франции, США и Китае) за 2012-2023 гг. [6-17].

Таблица 1

Объём производства автомобилей во Франции, США и Китае за 2012-2023гг. в миллионах единиц [6-17]

Порядковый номер года	Год	Франция	США	Китай
1	2012	1,967765	10,33263	19,27181
2	2013	1,74	11,0459	22,11683
3	2014	1,821464	11,6607	23,72289
4	2015	1,97	12,1001	24,50333
5	2016	2,082	12,19814	28,11879
6	2017	2,227	11,18999	29,01543
7	2018	2,2696	11,31471	27,8092
8	2019	2,20246	10,88002	25,72067
9	2020	1,316371	8,822399	25,22524
10	2021	1,351308	9,167214	26,08222
11	2022	1,383173	10,06034	27,02062
12	2023	1,505076	10,61156	30,16097

Особенностью рассматриваемого временного периода является то, что он включает годы мировой пандемии. Поэтому возникает вопрос, каким образом учитывать данные 2020-2021 годов в исследованиях. Будем искать значение коэффициента V значимости данных для периода 2021–2022 гг., решая задачу минимизации по V отклонения реального объема выпуска автомобилей за 2023 г. (по данным табл.1) от объема выпуска, рассчитанного по приведенным выше формулам. При этом для линейной модели сглаживания данных параметры a и b в (4) рассчитываются по данным за 2012–2022 гг., а коэффициенты значимости данных за все годы, кроме 2020–2021 гг. берутся равными единице. Для этого полагаем, что в задаче оптимизации (2) для конкретной страны число статистических данных по выпуску автомобилей составляет $n = 11$, значения x_i равны порядковому номеру года, значения y_i выбираются из соответствующего столбца табл.1, коэффициенты значимости данных $V_1 = V_2 = V_3 = ... = V_8 = V_{11} = 1$, а $V_9 = V_{10} = V$. Тогда оптимальное значение V, которое обозначим как V^*, ищется из решения задачи оптимизации

$$\min_{0 \le V \le 1} \delta(V) = \min_{0 \le V \le 1} \left| f(x_{12}, V) - y_{12} \right|. \quad (8)$$

В (8) $\delta(V)$ – это абсолютная погрешность (точность) оценки. Она равна отклонению расчетного значения выпуска автомобилей в конкретной стране от реального значения выпуска, если коэффициент значимости статистических данных в период пандемии равен V. Тогда $\delta(V^*)$ – это наименьшее значение указанной погрешности при оптимальном значении $V = V^*$. Следует подчеркнуть, что V^*=1 соответствует учету данных за годы пандемии в полном объеме, а значение V^*=0 означает исключение статистических данных за этот период из анализа.

В табл. 2 приведены $\delta(V)$ - точности оценки выпуска автомобилей в Китае за 2023 год в зависимости от коэффициента V значимости данных за годы пандемии при использовании линейной модели, в которой параметры определяются выражениями (4).

Таблица 2

Значение (в млн. машин) точности оценки выпуска автомобилей в 2023 году в Китае на линейной модели в зависимости от коэффициента значимости V данных в годы пандемии

V	0	0,01	0,014	0,015	0,016	0,1	0,5	1
$\delta(V)$	0,038	0,012	0,002	0,001	0,003	0,204	0,923	1,493

Представленные в табл. 2 результаты иллюстрируют тот факт, что оптимальное значение коэффициента V значимости статистических данных по выпуску автомобилей в Китае в период пандемии составляет $V^* = 0,015$. При этом значении V^* параметры линейной зависимости равны (после округления до тысячных): $a = 0,725$ и $b = 21,460$.

Таблица 3

Значение (в млн. машин) точности оценки выпуска автомобилей в 2023 году в США на линейной модели в зависимости от коэффициента значимости V данных в годы пандемии

V	0	0,06	0,07	0,08	0,1	0,2	0,5	1
$\delta(V)$	0,121	0,017	0,008	0,015	0,047	0,195	0,548	0,944

В табл. 3 приведены $\delta(V)$ - точности оценки выпуска автомобилей в США за 2023 г. в зависимости от коэффициента V значимости данных за годы пандемии при использовании линейной модели.

Табл. 3 иллюстрирует тот факт, что оптимальное значение коэффициента V значимости статистических данных по выпуску автомобилей в США в период пандемии составляет $V^* = 0,07$.

Отметим, что низкие значения V^* для Китая и США обусловлены сильным влиянием пандемии на выпуск автомобилей в этих странах. Однако для разных стран расчетные значения V^* коэффициента значимости статистических данных в годы пандемии сильно различаются. Так, для Франции для линейной модели оптимальное значение V^* равно

1. Этот факт иллюстрирует табл. 4, где приведены точности оценки выпуска автомобилей во Франции за 2023 год в зависимости от коэффициента значимости V статистических данных за годы пандемии при использовании линейной модели сглаживания.

Таблица 4

Значение (в млн. машин) точности оценки выпуска автомобилей в 2023 году во Франции на линейной модели в зависимости от коэффициента значимости V данных в годы пандемии

V	0	0,1	0,2	0,3	0,5	0,7	0,9	1
$\delta(V)$	0,356	0,325	0,280	0,240	0,172	0,118	0,072	0,052

Таким образом, статистические данные по выпуску автомобилей во Франции в годы пандемии лучше согласуются с линейной моделью их изменения, чем в Китае и США.

При использовании квадратичной модели сглаживания статистических данных по выпуску автомобилей во Франции за 2012-2021 гг. оптимальное значение коэффициента значимости данных за годы пандемии составляет $V^* = 0,33$. Это иллюстрирует табл. 5, в которой представлены $\delta(V)$ - точности оценки выпуска автомобилей в 2022 г. в зависимости от коэффициента значимости статистических данных за годы пандемии для этой страны.

Таблица 5

Значение (в млн. машин) точности оценки выпуска автомобилей в 2022 году во Франции на квадратичной модели в зависимости от коэффициента значимости V данных в годы пандемии

V	0	0,2	0,32	0,33	0,34	0,5	0,8	1
$\delta(V)$	1,262	0,212	0,008	0,004	0,039	0,155	0,288	0,336

Приведенные результаты исследований позволяют сделать вывод, что рассмотренный подход к анализу статистических данных обеспечивает возможность учета влияния пандемии на объем производства автомобилей в странах и, как следствие, скорректировать соответствующие прогнозные оценки на будущие периоды.

Таким образом, в статье предложена модификация метода наименьших квадратов для сглаживания статистических данных экономических показателей, состоящая в учете их значимости, а также продемонстрировано ее применение для анализа статистических данных в период пандемии для сферы автомобилестроения.

Список литературы

1. Линник Ю.В. Метод наименьших квадратов и основы теории обработки наблюдений. М.: Физматгиз. 1962. 333 с.
2. Королев Ю.Г. Метод наименьших квадратов в социально-экономических исследованиях: монография. М.: Статистика, 1980. 112 с.

3. Катулев А.Н., Кузнецов В.Н., Малевинский М.Ф., Соломаха Г.М. Двумерный полиномиальный фильтр // Известия РАН. Автоматика и телемеханика. 2003. №9. С. 77–88.
4. Катулев А.Н., Северцев Н.А., Соломаха Г.М. Исследование операций и обеспечение безопасности: прикладные задачи. М.: Наука, 2006. 240 с.
5. Васильев Ф.П. Численные методы решения экстремальных задач. М.: Наука, 1988. 552 с.
6. 2012 Statistics | www.oica.net // International Organization of Motor Vehicle Manufacturers. [Электронный ресурс] – URL: https://www.oica.net/category/production-statistics/2012-statistics/ – (Дата обращения 18.10.2024).
7. 2013 Statistics | www.oica.net // International Organization of Motor Vehicle Manufacturers. [Электронный ресурс] – URL: https://www.oica.net/category/production-statistics/2013-statistics/ – (Дата обращения 18.10.2024).
8. 2014 Statistics | www.oica.net // International Organization of Motor Vehicle Manufacturers. [Электронный ресурс] – URL: https://www.oica.net/category/production-statistics/2014-statistics/ – (Дата обращения 18.10.2024).
9. 2015 Statistics | www.oica.net // International Organization of Motor Vehicle Manufacturers. [Электронный ресурс] – URL: https://www.oica.net/category/production-statistics/2015-statistics/ – (Дата обращения 18.10.2024).
10. 2016 Statistics | www.oica.net // International Organization of Motor Vehicle Manufacturers. [Электронный ресурс] – URL: https://www.oica.net/category/production-statistics/2016-statistics/ – (Дата обращения 18.10.2024).
11. 2017 Statistics | www.oica.net // International Organization of Motor Vehicle Manufacturers. [Электронный ресурс] – URL: https://www.oica.net/category/production-statistics/2017-statistics/ – (Дата обращения 18.10.2024).
12. 2018 Statistics | www.oica.net // International Organization of Motor Vehicle Manufacturers. [Электронный ресурс] – URL: https://www.oica.net/category/production-statistics/2018-statistics/ – (Дата обращения 18.10.2024).
13. 2019 Statistics | www.oica.net // International Organization of Motor Vehicle Manufacturers. [Электронный ресурс] – URL: https://www.oica.net/category/production-statistics/2019-statistics/ – (Дата обращения 18.10.2024).
14. 2020 Statistics | www.oica.net // International Organization of Motor Vehicle Manufacturers. – . [Электронный ресурс] – Электрон. дан. – [Б.м., 2024]. – URL: https://www.oica.net/category/production-statistics/2020-statistics/– (Дата обращения 18.10.2024).
15. 2021 Statistics | www.oica.net // International Organization of Motor Vehicle Manufacturers. [Электронный ресурс] – URL: https://www.oica.net/category/production-statistics/2021-statistics/ – (Дата обращения 18.10.2024).
16. 2022 Statistics | www.oica.net // International Organization of Motor Vehicle Manufacturers. [Электронный ресурс] – URL: https://www.oica.net/category/production-statistics/2022-statistics/ – (Дата обращения 18.10.2024).
17. 2023 Statistics | www.oica.net // International Organization of Motor Vehicle Manufacturers. [Электронный ресурс] – URL: https://www.oica.net/category/production-statistics/2023-statistics/ – (Дата обращения 18.10.2024).

Об авторах

СОЛОМАХА Геннадий Михайлович – доктор физико-математических наук, доцент, профессор кафедры математической статистики и системного анализа, ФГБОУ ВО «Тверской государственный университет» (170100, г. Тверь, ул. Желябова, д. 33), e-mail: solomakha.gm@tversu.ru, ORCID 0009-0006-1084-6044, SPIN-код: 3458-6688.

ХИЖНЯК Станислав Виталиевич – аспирант, ФГБОУ ВО «Тверской государственный университет» (170100, г. Тверь, ул. Желябова, д. 33), e-mail: stanislav.khizhnyak@gmail.com, ORCID 0009-0008-7305-4348, SPIN-код: 8434-7096.

SMOOTHING OF STATISTICAL DATA ON ECONOMIC INDICATORS USING SIGNIFICANCE COEFFICIENTS

G.M. Solomakha, S.V. Khizhnyak

FGBOU VO "Tver State University", Tver

The article is dedicated to examining issues related to the smoothing of statistical data on economic indicators. The aim of this work is to refine the model for calculating parameters of linear and nonlinear regression of economic indicators when these data have varying significance. To achieve this, significance coefficients for statistical data are proposed. This approach requires solving nonlinear optimization problems to determine regression parameters, the solutions for which are presented in the article. An example of using the proposed approach to analyze car production in countries for a time period including the pandemic years is given. The scientific novelty of the results lies in the fact that, unlike existing data smoothing models, the models presented in the article allow for consideration of varying significance in statistical data when analyzing and forecasting economic indicators.

Keywords: linear and nonlinear data smoothing, least squares method, statistical data smoothing.

About the authors:

SOLOMAHA Gennadij Mihajlovich – Doctor of Physical and Mathematical Sciences, Associate Professor, Professor of the Department of Mathematical Statistics and System Analysis, FGBOU VO "Tver State University" (170100, Tver, Zhelyabova St., 33), e-mail: solomakha.gm@tversu.ru, ORCID 0009-0006-1084-6044, SPIN code: 3458-6688.

HIZHNJAK Stanislav Vitalievich – postgraduate student, FGBOU VO "Tver State University" (170100, Tver, Zhelyabova St., 33) e-mail: stanislav.khizhnyak@gmail.com, ORCID 0009-0008-7305-4348, SPIN code: 8434-7096.

Статья поступила в редакцию 20.11.2024 г.
Статья подписана в печать 16.12.2024 г.