

УДК 519.711.3:519.237

ОЦЕНИВАНИЕ ЛИНЕЙНОЙ РЕГРЕССИИ НА ОСНОВЕ
ОБОБЩЕННОГО МЕТОДА НАИМЕНЬШИХ МОДУЛЕЙ

А. Н. Тырсин, Л. А. Соколов

¹ Научно-инженерный центр «Надёжность и ресурс больших систем машин» УрО РАН,
620049 г. Екатеринбург, ул. Студенческая, 54 а.² Челябинский государственный университет,
454021 г. Челябинск, ул. Бр. Кашириных, 129.

E-mails: at2001@yandex.ru, 88slam88@mail.ru

Описан обобщенный метод наименьших модулей устойчивого нахождения оценок параметров линейных регрессионных моделей по экспериментальным данным. Доказаны теоремы существования и нахождения решения, а также состоятельности оценок. Приведены результаты исследования оценок параметров регрессии с помощью метода Монте–Карло на тестовых данных.

Ключевые слова: обобщенный метод наименьших модулей, оценки, регрессия, состоятельность, экспериментальные данные.

Введение. При обработке результатов экспериментальных исследований часто решается задача построения линейной регрессионной модели

$$\mathbf{y} = \mathbf{X}\mathbf{a} + \boldsymbol{\xi}, \quad (1)$$

где $\mathbf{y} = (y_1, y_2, \dots, y_n)^\top$ — вектор зависимой переменной; \mathbf{X} — регрессионная матрица размера $n \times (m+1)$ значений независимых (объясняющих) переменных; $\boldsymbol{\xi} = (\xi_1, \xi_2, \dots, \xi_n)^\top$ — вектор случайных ошибок; $\mathbf{a} = (a_0, a_1, \dots, a_m)^\top$ — искомый вектор параметров.

Относительно (1) обычно предполагают, что она удовлетворяет следующим условиям классической линейной модели множественной регрессии [1]:

- 1°) столбцы $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(m)}$ — неслучайные векторы;
- 2°) $\forall i \mathbb{E}[\xi_i] = 0$ (математическое ожидание ошибок равно нулю);
- 3°) $\forall i \mathbb{E}[\xi_i^2] = \mathbb{D}[\xi_i] = \sigma_\xi^2$ (гомоскедастичность ошибок);
- 4°) $\forall i \neq k \mathbb{E}[\xi_i \xi_k] = 0$;
- 5°) матрица \mathbf{X} имеет ранг $\text{Rg}(\mathbf{X}) = m + 1 < n$ и имеет вид

$$\mathbf{X} = \begin{pmatrix} x_1^{(0)} & x_1^{(1)} & \dots & x_1^{(m)} \\ x_2^{(0)} & x_2^{(1)} & \dots & x_2^{(m)} \\ \dots & \dots & \dots & \dots \\ x_n^{(0)} & x_n^{(1)} & \dots & x_n^{(m)} \end{pmatrix} = \begin{pmatrix} 1 & x_1^{(1)} & \dots & x_1^{(m)} \\ 1 & x_2^{(1)} & \dots & x_2^{(m)} \\ \dots & \dots & \dots & \dots \\ 1 & x_n^{(1)} & \dots & x_n^{(m)} \end{pmatrix}.$$

В этих условиях в случае, когда случайные ошибки ξ_i распределены по нормальному закону, эффективным методом определения вектора \mathbf{a} является метод наименьших квадратов (МНК). Реальные наблюдения во многих задачах в целом соответствуют данной предпосылке. Однако в ряде случаев возникают ситуации, когда предположение о гауссовой модели случайных

Александр Николаевич Тырсин (д.т.н.), ведущий научный сотрудник. Лев Александрович Соколов, аспирант, каф. теории управления и оптимизации.

ошибок не выполняется. В качестве примера можно привести задачи технической диагностики, когда требуется оценивать параметры модели (1) в условиях стохастической неоднородности, вызванной зарождающимися дефектами.

Известно, что даже незначительные нарушения указанных предпосылок резко снижают эффективность МНК-оценок. Отметим неустойчивость процедуры МНК-оценивания к наличию в измерениях выбросов [2]. С целью обеспечения устойчивости оценок к отклонению случайных ошибок от гауссовой модели разработан ряд робастных и непараметрических статистических методов. К числу наиболее распространенных процедур устойчивого оценивания можно отнести метод наименьших модулей (МНМ) [3] и знаковый метод [4]. Они основаны на более общих предположениях относительно случайных ошибок: требуется лишь их независимость и равновероятность принятия положительных и отрицательных значений. В работе рассматривается обобщенный метод наименьших модулей (ОМНМ), менее чувствительный к невыполнению указанных предположений [5].

1. Постановка задачи и основные предположения. ОМНМ-оценки вектора параметров регрессии (1) ищут в виде [5]

$$\hat{\mathbf{a}}_n = \arg \min_{\mathbf{a} \in \mathbb{R}^{m+1}} \sum_{i=1}^n \rho(|y_i - \mathbf{x}_i \mathbf{a}|), \quad (2)$$

где $\mathbf{x}_i = (x_i^{(0)}, \dots, x_i^{(m)}) = (1, x_i^{(1)}, \dots, x_i^{(m)})$, $i = 1, 2, \dots, n$; $\rho(\cdot)$ — некоторая монотонно возрастающая, дважды непрерывно дифференцируемая на положительной полуоси функция, причём

$$\rho(0) = 0, \quad \forall u > 0 \quad 0 < \rho'(u) < \infty, \quad -\infty < \rho''(u) < 0. \quad (3)$$

Введём относительно функции распределения $F(x)$ вектора случайных ошибок $\boldsymbol{\xi}$ следующие условия:

- 6°) случайные ошибки ξ_i с равными вероятностями принимают положительные и отрицательные значения, т.е. $F(0) = 0,5$;
- 7°) $F(x)$ удовлетворяет условию Липшица в нуле, т.е. $\forall \varepsilon > 0 \quad \forall u_1 \neq u_2 : |u_1| < \varepsilon, |u_2| < \varepsilon \exists L > 0 : |F(u_1) - F(u_2)| < L |u_1 - u_2|$;
- 8°) $\forall \varepsilon > 0 \quad \forall u_1 \neq u_2 : |u_1| < \varepsilon, |u_2| < \varepsilon \exists M > 0 : |F(u_1) - F(u_2)| > M |u_1 - u_2|$.

Также введём несколько условий относительно матрицы \mathbf{X} и вектора \mathbf{y} :

- 9°) $\text{Rg}(\mathbf{X}, \mathbf{y}) = m + 2 < n$;
- 10°) $\max_{i \in [1; n]} \max_{j \in [0; m+1]} |x_i^{(j)}| < \infty$;
- 11°) $\frac{1}{n} \mathbf{X}^T \mathbf{X} \rightarrow \boldsymbol{\Sigma} > 0$.

Отметим, что мы не требуем одновременного выполнения всех приведённых условий; они используются по мере необходимости.

Поставим следующие задачи: установим существование и возможность нахождения ОМНМ-оценок и исследуем вопрос об их состоятельности.

2. Существование и нахождение ОМНМ-оценок. Поскольку функция по-

теперь $\rho(\cdot)$ в (2) является вогнутой, то функция

$$Q(\mathbf{a}) = \sum_{i=1}^n \rho(|y_i - \mathbf{x}_i \mathbf{a}|) \quad (4)$$

также будет вогнутой, и возникает известная проблема точного нахождения ее минимума [7]. Справедлива следующая теорема.

ТЕОРЕМА 1. Пусть имеется выборка наблюдений (\mathbf{x}_i, y_i) , $i = 1, 2, \dots, n$, такая, что справедливо условие 9° , и задана на \mathbb{R}^{m+1} функция (4), у которой для функции потерь $\rho(\cdot)$ выполнены соотношения (3).

Введём гиперплоскости $\Omega_i = \Omega(\mathbf{a}, \mathbf{x}_i, y_i)$ в виде уравнений

$$y_i - \mathbf{x}_i \mathbf{a} = 0, \quad i = 1, 2, \dots, n. \quad (5)$$

Введём также узловые точки пересечения гиперплоскостей (5):

$$u = \bigcap_{s \in M} \Omega_s, \quad M = \{k_1, \dots, k_{m+1}\}, \quad k_1 < k_2 < \dots < k_{m+1}, \quad k_l \in \{1, \dots, n\}. \quad (6)$$

Обозначим через U множество всех узловых точек (6). Тогда

- 1) для того чтобы точка $\mathbf{a}^\circ = (a_0^\circ, a_1^\circ, \dots, a_m^\circ)^\top$ являлась точкой локального минимума функции (4), необходимо, чтобы $\mathbf{a}^\circ \in U$;
- 2) функция (4) на \mathbb{R}^{m+1} имеет хотя бы одну точку локального минимума, причём число локальных минимумов функции (4) конечно.

Доказательство. Вначале докажем утверждение 1. Очевидно, что все гиперплоскости (5) являются особыми точками функции (4), поскольку в них она недифференцируема.

Рассмотрим произвольную стационарную точку $\mathbf{a}^\circ = (a_0^\circ, a_1^\circ, \dots, a_m^\circ)^\top$, в которой градиент $\nabla Q(\mathbf{a}^\circ) = 0$. Предположим, что \mathbf{a}° является точкой локального минимума. Поскольку эта точка не является особой, то функция (4) в ней дважды дифференцируема. Тогда матрица Гессе $\mathbf{H}(\mathbf{a}^\circ)$ должна быть неотрицательно определенной, т.е. все угловые миноры должны быть неотрицательными: $\Delta_1 \geq 0$, $\Delta_2 \geq 0$, \dots , $\Delta_{m+1} \geq 0$ [9]. Определим знак первого углового минора Δ_1 . В силу строгой вогнутости функции $\rho(\cdot)$ имеем для всех неособых точек

$$\Delta_1 = \frac{\partial^2 Q(\mathbf{a}^\circ)}{\partial (\mathbf{a}^\circ)^2} = \sum_{i=1}^n \rho'' \left(\left| y_i - a_0^\circ - \sum_{j=1}^m a_j^\circ x_{ij} \right| \right) < 0.$$

Получено противоречие. Таким образом, локальные минимумы, если они существуют, могут находиться только в особых точках.

Покажем, что локальные минимумы, если они существуют, могут находиться только в узловых точках (6), являющихся пересечением $(m+1)$ произвольно взятых взаимно непараллельных гиперплоскостей $\Omega_1, \Omega_2, \dots, \Omega_{m+1}$. Такие узловые точки существуют ввиду выполнения условия 9° .

Предположим, что некоторая точка $\mathbf{a}^k \in \mathbb{R}^{m+1}$ гиперплоскости Ω_k является локальным минимумом функции (4). Пусть эта точка не является

узловой. Следовательно, она образована пересечением не более чем m произвольно взятых взаимно непараллельных гиперплоскостей $\Omega_1, \Omega_2, \dots, \Omega_m$. Поэтому найдется такой ненулевой вектор $\mathbf{z}^k \in \mathbb{R}^{m+1}$, который будет параллелен всем гиперплоскостям $\Omega_1, \Omega_2, \dots, \Omega_m$ [10]. Очевидно, что производная функции (4) в точке \mathbf{a}^k по направлению \mathbf{z}^k существует. Поэтому с учётом доказанного выше точка \mathbf{a}^k не является локальным минимумом функции (4) по направлению \mathbf{z}^k . Получено противоречие. Тогда локальные минимумы функции (4), если они существуют, могут быть только в узловых точках.

Докажем теперь утверждение 2. Как было показано выше, локальные минимумы могут быть только в узловых точках, число которых конечно и не превосходит C_n^{m+1} . Поэтому число локальных минимумов конечно и ограничено числом узловых точек.

Покажем теперь, что множество локальных минимумов не пусто.

Поскольку функция (4) является непрерывной и ограниченной снизу, то она имеет точную нижнюю грань, которая достигается хотя бы в одной из узловых точек. Покажем это.

Предположим противное. Пусть нижняя грань функции (4) достигается в точке $\mathbf{a}^0 \notin U$. Поскольку точка \mathbf{a}^0 не является точкой локального минимума, следовательно, найдётся некоторая точка \mathbf{a} , такая, что $Q(\mathbf{a}) < Q(\mathbf{a}^0)$. Это означает, что нижняя грань функции (4) не может достигаться в точке \mathbf{a}^0 . Получено противоречие. Следовательно, точная нижняя грань функции (4) достигается на множестве узловых точек U .

Пусть точная нижняя грань достигается в некоторой узловой точке \mathbf{a}^* . Рассмотрим произвольную точку \mathbf{a} , лежащую в сколь угодно малой δ -окрестности точки \mathbf{a}^* , в которую не попадает ни одна из остальных узловых точек. По определению нижней грани $Q(\mathbf{a}) \geq Q(\mathbf{a}^*)$. Поскольку при этом точка \mathbf{a} не является узловой, она не может быть локальным минимумом. Значит, найдётся некоторая точка \mathbf{a}^1 , такая, что $Q(\mathbf{a}) > Q(\mathbf{a}^1)$. Но, с другой стороны, $Q(\mathbf{a}^1) \geq Q(\mathbf{a}^*)$. Отсюда $Q(\mathbf{a}) > Q(\mathbf{a}^1) \geq Q(\mathbf{a}^*)$, то есть \mathbf{a}^* — локальный минимум. Теорема [1] доказана. \square

3. Состоятельность ОМНМ-оценок. В [8] показано, что ОМНМ можно свести к взвешенному методу наименьших модулей (ВМНМ), т. е. вместо (2) решать задачу

$$W(\mathbf{a}) = \sum_{i=1}^n p_i |y_i - \mathbf{x}_i \mathbf{a}| \rightarrow \min_{\mathbf{a} \in \mathbb{R}^{m+1}}, \quad (7)$$

где p_i — весовые коэффициенты, причём $\forall i p_i > 0$. Если $|y_i - \mathbf{x}_i \mathbf{a}| > 0$, то $p_i = \rho(|y_i - \mathbf{x}_i \mathbf{a}|) / |y_i - \mathbf{x}_i \mathbf{a}| > 0$, а если $|y_i - \mathbf{x}_i \mathbf{a}| = 0$, то в качестве p_i можно взять любое положительное число, например 1. Поэтому для доказательства состоятельности ОМНМ-оценок достаточно доказать состоятельность ВМНМ-оценок для произвольных положительных весовых коэффициентов.

Очевидно, что минимум функции (7) удовлетворяет соотношению

$$\hat{\mathbf{a}}_n = \arg \min_{\mathbf{a} \in \mathbb{R}^{m+1}} \sum_{j=0}^m \left(\sum_{i=1}^n x_i^{(j)} p_i \operatorname{sign}(y_i - \mathbf{x}_i \mathbf{a}) \right)^2, \quad (8)$$

которое отличается от известной знаковой задачи [4, с. 69] множителями p_i у слагаемых под знаком суммы. Это делает естественным доказательство

состоятельности ВМНМ-оценок по схеме, предложенной в [4]. Справедлива следующая теорема.

ТЕОРЕМА 2. *Предположим, что в (1) случайные ошибки $\xi_1, \xi_2, \dots, \xi_n$ независимы, одинаково распределены и их общая функция распределения $F(x)$ удовлетворяет условиям $6^\circ, 7^\circ, 8^\circ$. Пусть для матрицы \mathbf{X} и вектора \mathbf{y} справедливы условия $9^\circ, 10^\circ, 11^\circ$. Тогда ВМНМ-оценка $\hat{\mathbf{a}}_n$ состоятельна.*

Доказательство. Доказательство выполним по схеме, приведённой в [4, с. 81–85] для доказательства состоятельности знаковых оценок параметров линейной регрессии. Обозначим через \mathbf{a}° вектор истинных параметров регрессии. Умножим целевую функцию ВМНМ-оценки (8) на $1/n$ и сделаем замену переменных, положив $\mathbf{a} = \mathbf{a}^\circ + \mathbf{t}$. Тогда $\hat{\mathbf{a}}_n = \mathbf{a}^\circ + \hat{\mathbf{t}}_n$, где

$$\hat{\mathbf{t}}_n = \arg \min_{\mathbf{t} \in \mathbb{R}^{m+1}} \frac{1}{n} \sum_{j=0}^m \left(\sum_{i=1}^n x_i^{(j)} p_i \operatorname{sign}(\mathbf{x}_i \mathbf{t} - \xi_i) \right)^2. \quad (9)$$

Остаётся показать, что $\hat{\mathbf{t}}_n \xrightarrow{\mathbf{P}} 0$. Доказательство разобьём на две части.

Часть 1. Начнём с утверждения, что для каждого $j = 0, 1, \dots, m$ и любого $T > 0$

$$\sup_{|\mathbf{t}| \leq T} \left| \frac{1}{n} \sum_{i=1}^n x_i^{(j)} p_i \operatorname{sign}(\mathbf{x}_i \mathbf{t} - \xi_i) - \mathbb{E} \left[\frac{1}{n} \sum_{i=1}^n x_i^{(j)} p_i \operatorname{sign}(\mathbf{x}_i \mathbf{t} - \xi_i) \right] \right| \xrightarrow{\mathbf{P}} 0.$$

В силу равномерного закона больших чисел [4, с. 90–94]

$$\arg \min_{|\mathbf{t}| \leq T} \left| \frac{1}{n} \sum_{i=1}^n x_i^{(j)} p_i \operatorname{sign}(\mathbf{x}_i \mathbf{t} - \xi_i) \right|^2 \quad (10)$$

при $n \rightarrow \infty$ сближается (по вероятности) с

$$\arg \min_{|\mathbf{t}| \leq T} \left| \frac{1}{n} \mathbb{E} \left[\sum_{i=1}^n x_i^{(j)} p_i \operatorname{sign}(\mathbf{x}_i \mathbf{t} - \xi_i) \right] \right|^2$$

при любом значении $T > 0$. С другой стороны,

$$\mathbb{E} \left[\frac{1}{n} \sum_{i=1}^n x_i^{(j)} p_i \operatorname{sign}(\mathbf{x}_i \mathbf{t} - \xi_i) \right] = \frac{1}{n} \sum_{i=1}^n x_i^{(j)} p_i (1 - 2F(\mathbf{x}_i \mathbf{t})). \quad (11)$$

Далее, поскольку $F(0) = 0,5$,

$$\arg \min_{|\mathbf{t}| \leq T} \left| \frac{1}{n} \mathbb{E} \left[\sum_{i=1}^n x_i^{(j)} p_i \operatorname{sign}(\mathbf{x}_i \mathbf{t} - \xi_i) \right] \right|^2 = 0,$$

причём для достаточно больших n точка минимума $\mathbf{t} = 0$ является единственной. Покажем это.

Предположим противное. Пусть существует $\mathbf{t} \neq 0$, для которого функция (11) равна нулю при каждом $j = 0, 1, \dots, m$. Умножив правую часть выражения (11) на t_j и просуммировав по j , получим

$$\frac{1}{n} \sum_{i=1}^n (\mathbf{x}_i \mathbf{t}) p_i (1 - 2F(\mathbf{x}_i \mathbf{t})) = 0. \quad (12)$$

Поскольку все $p_i > 0$, то легко видеть, что $\forall u \in \mathbb{R}^1 u p_i (1 - 2F(u)) \leq 0$, причём равенство нулю при выполнении условия 8° достигается только при $u = 0$. Поэтому равенство (12) справедливо только в том случае, когда $(\mathbf{x}_i \mathbf{t}) = 0$, $i = 1, 2, \dots, n$. Тогда $(\sum_{i=1}^n \mathbf{x}_i) \mathbf{t} = 0$, и, следовательно, $\mathbf{t}^\top (\frac{1}{n} \mathbf{X}^\top \mathbf{X}) \mathbf{t} = 0$ для данного $\mathbf{t} \neq 0$. Отсюда $\frac{1}{n} \mathbf{X}^\top \mathbf{X} = 0$, что противоречит условию 11°. Таким образом, (10) сходится к нулю по вероятности при $n \rightarrow \infty$ и любом $T > 0$. Отсюда следует, что для любого $T > 0$

$$\arg \min_{\mathbf{t} \in \mathbb{R}^{m+1}} \frac{1}{n} \sum_{j=0}^m \left(\sum_{i=1}^n x_i^{(j)} p_i \operatorname{sign}(\mathbf{x}_i \mathbf{t} - \xi_i) \right)^2 \xrightarrow{\mathbf{P}} 0.$$

Часть 2. Покажем, что существуют такие $b > 0$ и $T > 0$, что с вероятностью, сколь угодно близкой к 1,

$$\sum_{j=0}^m \left(\frac{1}{n} \sum_{i=1}^n x_i^{(j)} p_i \operatorname{sign}(\mathbf{x}_i \mathbf{t} - \xi_i) \right)^2 > b \quad (13)$$

для достаточно больших n , если $|\mathbf{t}| > T$. Тем самым будет доказано, что минимум выражения (9) при достаточно больших n с вероятностью, сколь угодно близкой к 1, лежит внутри компакта $|\mathbf{t}| \leq T$. Вместе с результатом части 1 неравенство (13) доказывает состоятельность ВМНМ-оценки.

Докажем (13). Рассмотрим вектор $\varphi(\mathbf{t}) = (\varphi_0(\mathbf{t}), \varphi_1(\mathbf{t}), \dots, \varphi_m(\mathbf{t}))$, где

$$\varphi_j(\mathbf{t}) = \frac{1}{n} \sum_{i=1}^n x_i^{(j)} p_i \operatorname{sign}(\mathbf{x}_i \mathbf{t} - \xi_i).$$

Введём единичный вектор $\mathbf{t}^0 = \mathbf{t}/|\mathbf{t}|$ и покажем, что существуют $b > 0$ и $T > 0$ такие, что для достаточно больших n с вероятностью, сколь угодно близкой к 1, для всех $|\mathbf{t}| \geq T$ выполняется неравенство $|\varphi(\mathbf{t}) \mathbf{t}^0|^2 > b$, т. е. получим, что

$$\frac{1}{|\mathbf{t}|^2} \left(\frac{1}{n} \sum_{i=1}^n (\mathbf{x}_i \mathbf{t}) p_i \operatorname{sign}(\mathbf{x}_i \mathbf{t} - \xi_i) \right)^2 > b. \quad (14)$$

Из неравенства (14) и неравенства Коши—Шварца будет следовать (13). Заметим, что если $|\mathbf{x}_i \mathbf{t}| > |\xi_i|$, то $(\mathbf{x}_i \mathbf{t}) p_i \operatorname{sign}(\mathbf{x}_i \mathbf{t} - \xi_i) = p_i |\mathbf{x}_i \mathbf{t}|$.

Воспользовавшись тем, что $|A + B| \geq |A| - |B|$, запишем

$$\left| \frac{1}{n} \sum_{i=1}^n (\mathbf{x}_i \mathbf{t}) p_i \operatorname{sign}(\mathbf{x}_i \mathbf{t} - \xi_i) \right| \geq \frac{1}{n} \sum_{i=1}^n |(\mathbf{x}_i \mathbf{t}) p_i| - \frac{2}{n} \sum_{\{i: |\mathbf{x}_i \mathbf{t}| \leq |\xi_i|\}} |(\mathbf{x}_i \mathbf{t}) p_i|. \quad (15)$$

Оценим сверху второе слагаемое в правой части последнего выражения. Заметим, что $\forall R > 0$

$$|(\mathbf{x}_i \mathbf{t}) p_i| \mathbf{I}(|(\mathbf{x}_i \mathbf{t}) p_i| \leq |\xi_i|) \leq R \mathbf{I}(|\xi_i| \leq R) + |(\mathbf{x}_i \mathbf{t}) p_i| \mathbf{I}(|\xi_i| > R).$$

Выберем R так, чтобы $1 - 2\mathbf{P}\{|\xi_i| > R\} > 0$. В результате для второго слагаемого в правой части (15) получаем оценку

$$\frac{2}{n} \sum_{\{i: |\mathbf{x}_i \mathbf{t}| \leq |\xi_i|\}} |(\mathbf{x}_i \mathbf{t}) p_i| \leq \frac{2}{n} \sum_{i=1}^n \left\{ R \mathbf{I}(|\xi_i| \leq R) + |(\mathbf{x}_i \mathbf{t}) p_i| \mathbf{I}(|\xi_i| > R) \right\}.$$

В силу условия 10° и закона больших чисел при $n \rightarrow \infty$ имеем

$$2R\mathbf{P}\{|\xi_i| \leq R\} + 2\mathbf{P}\{|\xi_i| > R\} \frac{1}{n} \sum_{i=1}^n |(\mathbf{x}_i \mathbf{t}) p_i| + o_p(1).$$

Тогда, продолжив (15), получим, что для любого $\varepsilon > 0$ с вероятностью, сколь угодно близкой к 1, для достаточно больших n

$$\begin{aligned} \left| \frac{1}{n} \sum_{i=1}^n (\mathbf{x}_i \mathbf{t}) p_i \operatorname{sign}(\mathbf{x}_i \mathbf{t} - \xi_i) \right| &\geq \\ &\geq -2R\mathbf{P}\{|\xi_i| \leq R\} + (1 - 2\mathbf{P}\{|\xi_i| > R\}) \frac{1}{n} \sum_{i=1}^n |(\mathbf{x}_i \mathbf{t}) p_i| - \varepsilon. \end{aligned} \quad (16)$$

При условиях 10° и 11° имеем, что существует $K > 0$, для которого

$$\frac{1}{n} \sum_{i=1}^n |(\mathbf{x}_i \mathbf{t}) p_i|^2 < K \left(\frac{1}{n} \sum_{i=1}^n |(\mathbf{x}_i \mathbf{t}) p_i| \right) |\mathbf{t}|, \quad (17)$$

$$\frac{1}{n} \sum_{i=1}^n |(\mathbf{x}_i \mathbf{t}) p_i|^2 = \frac{1}{n} \sum_{i=1}^n ((\mathbf{x}_i \mathbf{t}) p_i)^\top ((\mathbf{x}_i \mathbf{t}) p_i) = \mathbf{t}^\top (\mathbf{X}_p^\top \mathbf{X}) \mathbf{t} > k |\mathbf{t}|^2,$$

где k — некоторое положительное число, $\mathbf{X}_p = \{x_i^{(j)} p_i\}$. Тогда при больших n

$$\frac{1}{n} \sum_{i=1}^n |(\mathbf{x}_i \mathbf{t}) p_i|^2 > \frac{k}{2} |\mathbf{t}|^2. \quad (18)$$

Из (17) и (18) следует, что для достаточно больших n

$$\frac{1}{n} \sum_{i=1}^n |(\mathbf{x}_i \mathbf{t}) p_i| > \frac{k}{2K} |\mathbf{t}|^2. \quad (19)$$

Вернувшись к (16), получаем с учётом (19), что при достаточно больших n со сколь угодно близкой к 1 вероятностью

$$\left| \frac{1}{n} \sum_{i=1}^n (\mathbf{x}_i \mathbf{t}) p_i \operatorname{sign}(\mathbf{x}_i \mathbf{t} - \xi_i) \right| > \tilde{k} |\mathbf{t}| - 2R\mathbf{P}\{|\xi_i| \leq R\} - \varepsilon,$$

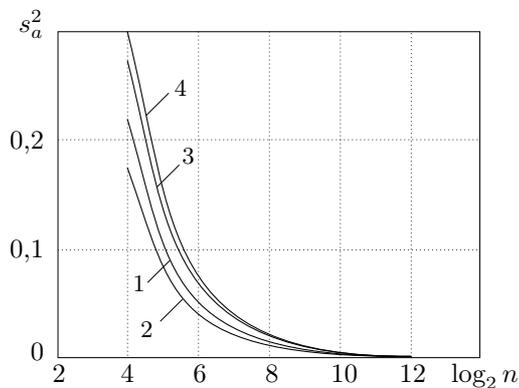
где $\tilde{k} = \frac{k}{2K} (1 - 2P\{|\xi_i| > R\})$. Вернувшись к (19), получаем, что при достаточно больших $|\mathbf{t}|$ и достаточно больших n

$$|\varphi(\mathbf{t})\mathbf{t}^0|^2 > \frac{(\tilde{k}|\mathbf{t}| - 2RP\{|\xi_i| \leq R\})^2}{|\mathbf{t}|^2} > b > 0$$

со сколь угодно близкой к 1 вероятностью, что и требовалось доказать. Из состоятельности ВМНМ-оценки (8) для произвольных положительных весовых коэффициентов p_i , следует состоятельность ОМНМ-оценки (2). \square

4. Статистическое моделирование методом Монте—Карло. Поскольку при доказательстве состоятельности не оценивается точность приближения $\hat{\mathbf{a}}_n \approx \mathbf{a}$ при конечном числе наблюдений n , представляет интерес такая оценка. Выполним такую оценку на основе метода Монте—Карло, который позволяет контролировать точность оценивания. Результаты моделирования показали, что при увеличении n точность ОМНМ-оценок сопоставима с точностью МНК-оценок. В качестве иллюстрации приведём следующий пример.

ПРИМЕР. Исследуем методом Монте—Карло оценку \hat{a}_n параметра a в однопараметрической модели $y_i = a + \xi_i$, $i = 1, 2, \dots, n$. Зададим для случайных ошибок ξ_i распределение с засорением и смещением в виде смеси двух нормальных распределений $0,75N(0; 1) + 0,25N(2; 3^2)$. Для серии из 5000 повторений с ростом объёма выборки n получаем, что дисперсия ОМНМ-оценки монотонно убывает (см. рисунок). Скорость сходимости также практически не уступает скорости сходимости оценок наименьших квадратов и модулей.



Зависимость дисперсии оценок от объёма выборки ($\xi_i \sim 0,75N(0; 1) + 0,25N(2; 3^2)$): 1 — МНК-оценка; 2 — МНМ-оценка; 3 — ОМНМ-оценка для $\rho(x) = |x|^{0,5}$; 4 — ОМНМ-оценка для $\rho(x) = \arctg|x|$

Замечание. Теорема 2 доказана для случая $F(0) = 0,5$. Вопрос о состоятельности ОМНМ-оценок для случая $F(0) \neq 0,5$ остается открытым, однако результаты статистического моделирования свидетельствуют о состоятельности ОМНМ-оценок и в этом случае.

5. Выводы.

1. Доказаны теоремы существования и нахождения ОМНМ-оценок, причём в отличие от МНМ-оценок число решений конечно и может быть оценено.
2. Установлено, что ОМНМ-оценки параметров регрессии состоятельны в случае, когда случайные ошибки с равными вероятностями принимают положительные и отрицательные значения и выполнен ряд других ограничений.
3. Результаты исследования ОМНМ-оценок с помощью метода Монте—Карло на тестовых данных, в которых случайные ошибки имеют различное засорение, свидетельствуют о состоятельности ОМНМ-оценок и в более общем случае. В частности, не требуется, чтобы случайные ошибки с равными вероятностями принимали положительные и отрицательные значения.

Работа выполнена при поддержке РФФИ (проект № 10-01-96013-р-урал-а).

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Королюк В. С., Портенко Н. И., Скороход А. В., Турбин А. Ф.* Справочник по теории вероятностей и математической статистике. — М.: Наука, 1985. — 640 с.
2. *Хьюбер П.* Робастность в статистике. — М.: Мир, 1984. — 303 с.
3. *Мудров В. И., Кушко В. Л.* Методы обработки измерений. Квазиправдоподобные оценки. — М.: Радио и связь, 1983. — 304 с.
4. *Болдин М. В., Симонова Г. И., Тюрин Ю. Н.* Знаковый статистический анализ линейных моделей. — М.: Наука, 1997. — 226 с.
5. *Тырсин А. Н.* Робастное построение регрессионных зависимостей на основе обобщенного метода наименьших модулей / В сб.: *Вероятность и статистика. 9 /* Зап. научн. сем. ПОМИ. — СПб.: ПОМИ, 2005. — Т. 328. — С. 236–250; англ. пер.: *Tyrstin A. N.* Robust construction of regression models based on the generalized least absolute deviations method // *J. Math. Sci. (N. Y.)*, 2006. — Vol. 139, No. 3. — P. 6634–6642.
6. *Тырсин А. Н.* Исследование свойств обобщенного метода наименьших модулей (на примере оценки параметра сдвига) // *Заводская лаборатория*, 2007. — Т. 73, № 11. — С. 71–76.
7. *Рокафеллар Р.* Выпуклый анализ. — М.: Мир, 1973. — 472 с.
8. *Панюков А. В., Тырсин А. Н.* Взаимосвязь взвешенного и обобщенного вариантов метода наименьших модулей // *Извест. Челябинск. научн. центра*, 2007. — № 1(35). — С. 6–11.
9. *Галеев Э. М.* Оптимизация: теория, примеры, задачи. — М.: Едиториал УРСС, 2002. — 302 с.
10. *Ефимов Н. В., Розендорн Э. Р.* Линейная алгебра и многомерная геометрия. — М.: Физматлит, 1970. — 400 с.

Поступила в редакцию 01/II/2010;
в окончательном варианте — 10/X/2010.

MSC: 60-99, 62-99

THE ESTIMATION OF LINEAR REGRESSION IS BASED ON THE GENERALIZED LEAST MODULES METHOD

A. N. Tyrstin, L. A. Sokolov

¹ Reliability and Resource of Large Machine and Systems, Science and Engineering Center, Ural Branch of RAS, 54 a, Studencheskaya st., Ekaterinburg, 620049, Russia.

² Chelyabinsk State University, 129, Br. Kashirinih st., Chelyabinsk, 454021, Russia.

E-mails: at2001@yandex.ru, 88slam88@mail.ru

The generalized least modules method is shown in this paper. It can be applied to find estimations of parameters of the linear regression model that is based on experimental data. The theorems of existence and finding of solution are proved. The consistency of estimator is proved as well. The results of investigation of regression parameters are demonstrated here. Monte–Carlo method was used for this investigation.

Key words: *generalized least modules method, estimations, regression, consistency, experimental data.*

Original article submitted 01/II/2010;
revision submitted 10/X/2010.

Alexandr N. Tyrstin (Dr. Sci. (Techn.)), Leading Research Scientist. *Lev A. Sokolov*, Postgraduate Student, Dept. of Control & Optimization Theory.